Award Number: W81XWH-06-2-0057

TITLE: Advanced Technologies in Safe and Efficient Operating Rooms

PRINCIPAL INVESTIGATOR: Adrian E. Park, MD

CONTRACTING ORGANIZATION: University of Maryland Medical Center
Baltimore, Maryland

REPORT DATE: Qevqdgt''2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-10-2009	Annual	26 SEP 2008 – 25 SEP 2009

4. TITLE AND SUBTITLE
Advanced Technologies in Safe and Efficient Operating Rooms

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-2-0057

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Adrian E. Park, MD

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

E-Mail:

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Maryland Medical Center
Baltimore, MD 21201

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
No abstract provided.

15. SUBJECT TERMS
No subject terms provided.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	141	

Table of Contents

.

1. Current staff role and percentage effort on each project:

STAFF MEMBER	Role	% EFFORT
A.Park, MD	PI	15
B. Jarrell, MD	Co-PI	5
R. Shekhar, Ph.D	Researcher	25
G. Moses, Ph.D	Director	40
P. Nagy, Ph.D	Researcher	20
S. Kavic, MD	Clinical	10
P. Turner	Clinical	10

2. Contract expenditures to date (as applicable):

COST ELEMENTS	THIS QUARTER	CUMULATIVE
Personnel		411,856
Fringe Benefits		67,780
Supplies		-6337
Equipment		0
Travel		1,665
Other Direct Costs		889,862
Subtotal		1,312,143
Indirect Costs		154,277
Fee		
Total		1,519,103

Introduction

During the past decade, we witnessed an extraordinary evolution in surgical care based upon rapid advances in technology and creative approaches to medicine. The increased speed and power of computer applications, the rise of visualization technologies related to imaging and image guidance, improvement in simulation-based technologies (tissue properties, tool-tissue interaction, graphics, haptics, etc) has caused an explosion in surgical advances. That said, we remain far behind scientists in applying information systems to patient care. This research effort has proceeded under the mantle of "Operating Room of the Future" research. We replaced that theme with the more appropriate "Innovations in the Surgical Environment."

The content of this annual report contains information pertinent to continued activities in relation to the W81XWH-06-2-0057, "Advanced Technologies in Safe and Efficient Operating Rooms" project. This contract consists of a scope of work that fits seamlessly onto a prior research activity in the contract DAMD-17-03-2-0001, "Advanced technologies in safe and efficient operating rooms" work. The current research project activities are based upon three pillars of research, OR Informatics, Simulation for Training and Smart Image. A fourth research area was included in the Informatics pillar during this period of performance that targeted physical and cognitive ergonomics/human factors.

Two of the Informatics projects were closed during this year. The Intra Perioperative Communication (IPC) project has been completed; the CAST project was replaced by the Video Summarization project. In the Simulation pillar, the Maryland Virtual Patient (MVP) project has been concluded other than for preparation of manuscripts and presentations. Other sources of funding for this project are being sought. Work continues on the other projects under the terms of a no-cost extension. Milestones and termination dates for these projects were projected and reported.

Body

A. OR Informatics

Informatics subgroup 1. Workflow and Operations Research for Quality (WORQ)

The Perioperative Scheduling Study is looking at how using post-operative destination information during the process of surgery scheduling can influence congestion in postoperative units such as ICUs and IMCs, which lead to overnight boarders in the PACU. The research team is Jeffrey W. Herrmann, Ph.D., and Greg Brown, a graduate student, both with the University of Maryland, College Park. The team is working closely with Michael Harrington, Ramon Konewko, R.N., and Paul Nagy, Ph.D., for guidance and assistance. This research is summarized in a Ph.D. Preliminary Oral Exam, entitled The Surgery Scheduling Problem, Block Release Policies, and Operations Research Applied to Health Care; by William Herring under the mentorship of Dr. Hermann. **The slides for this presentation are placed in the Appendices to this report.**

4

We have developed a mathematical evaluation model for evaluating congestion in post-operative units, including ICUs, IMCs, and floor units. This model requires data about post-operative destinations and length-of-stay distributions for different types of surgeries. We have analyzed data about cardiac surgeries from two years and have analyzed UMMC financial records for all of the surgical cases for fiscal year 2007. We developed an algorithm for predicting bed requirements based on the surgical schedule and have conducted a preliminary study comparing these predictions to other prediction methods for two units. The preliminary results show that the new bed requirements prediction method is more accurate. We plan to complete the study and document the results in a technical report this fall. We continue to refine and implement mathematical models for evaluating how different block release policies affect OR utilization and staff overtime.

A summary of doctoral level work performed by William Herring in support of this project is included here. During the Spring 2009 semester, I conducted a thorough review of the operations research literature on operating room (OR) scheduling. In the course of this review I came across what I believe to be a critical and understudied interaction that has been the focus of my research since then. For many hospitals, including UMMC, the initial stage of the surgery scheduling process is the allocation of available operating room blocks to different surgical service lines. However, as the schedule for a given day evolves, focus shifts to individual patients and a new set of challenges present themselves to operating room managers.

A great deal of research has been conducted on algorithms for scheduling individual patients into available operating room space, and in recent years a good deal of attention has been paid to determining the best ways to allocate operating room blocks. However, very little work has been done on the interaction between these two pieces of the scheduling puzzle. In order to systematically explore the policies that affect this interaction, I worked closely with members of UMMC's perioperative staff to observe all stages of the scheduling process and develop a model for how the process evolves as the day of surgery approaches. In developing this model, I determined that the key policies that control this interaction are the block release policy (when OR managers take unused space back from individual service lines) and the request queue placement policies (how this space is used).

In August, I developed a stochastic dynamic programming (SDP) formulation of the single day surgery scheduling problem which incorporates the block schedule and allows for flexibility in setting and testing the effectiveness of different block release and request queue policies. A key component of the formulation is the arrival process for the demand for OR space (both the quantity *and* the timing of the demand). In order to estimate this demand, I have been working to get access to data from UMMC's CDR and went through a training course on pulling data tables from the Clinical Data repository (CDR). Also, because the optimal policy suggested by the SDP formulation might not be practical from a OR manager's perspective, I developed a simpler decision-making model which I feel more closely reflects how request queue decisions are currently made.

In September, I wrote a computer program that solves small instances of the SDP and began exploring the types of policies that the model suggests. In order to compare the optimal policies suggested by the SDP with more practical policies, the program is flexible enough to accept policy constraints and only produce solutions that operate within those constraints. Since I do not yet have accurate estimates of the demand for surgery, these initial runs are being done with simple demand distributions and I am testing the formulation's sensitivity to different types of distributions. I expect this work to continue in the following months, as I incorporate real data from the CDR and attempt to solve larger, more realistic versions of the model.

In August, I attended the Mayo Clinic Conference on Systems Engineering and Operations Research in Health Care. As a student just beginning my research in health care operations, this conference served as my introduction to formal research conferences. The goals of the conference generally fell into three categories: (1) exposure to new problems and methods being explored in this research area, (2) the opportunity to network with similarly-minded professionals, and (3) dialogue on the challenges of implementing research findings in complex hospital and clinical environments. While all conferences can be expected to meet the first two of these goals, what truly made the conference meaningful was its success in bringing together physicians, administrators, and researchers to address the third goal.

As mentioned above, the first chapter of my dissertation seeks to develop a model for the surgery scheduling process for a large operating room (OR) suite, a problem which involves decision-making in a highly stochastic environment. A typical modeling approach for this type of problem is to use what is known as a Markov decision process (MDP), and one of the presenters gave a talk on using an MDP to analyze scheduling decisions. Several other talks presented simulation models applied to scheduling. Both of these methodologies can be applied to my problem, and it was useful and encouraging to see them applied successfully to similar problems.

I was also fortunate enough while in Rochester to meet and spend a couple hours with the Operations Manager for the Mayo Clinic's Department of Surgery, and he generously shared with me many of the details of their scheduling system. This networking opportunity was a unique learning experience as I strive to make my research general enough to apply in a wide range of settings. Finally, several talks at the conference were directed at bridging the gap between systems engineers/operations researchers (SE/OR) and health care providers as the health care system moves to incorporate more evidence-based practices into its operations. Naturally, health care providers approach operational decisions with individual patient care at the forefront of their thinking, while SE/OR practitioners are trained to look at the same problems from a system-wide perspective. This difference in perspective creates the need for careful communication and collaboration between the different stakeholders. By bringing together physicians, administrators, and researchers under the same roof, this annual conference facilitates this communication and assists the effective implementation of SE/OR based practices across the health care system.

Informatics subgroup 2. Operating Room Glitch Analysis (OGA)

The OGA project, focusing on institutional learning, examined the workflow around performance indicators in the perioperative environment and building a graphical dashboard to allow data mining and trend analysis of operating indicators.

The dashboard was constructed using the Ruby on Rails web development platform with a MySQL database dynamically driving the queries. An interactive graphical dashboard provided synthesis around delays in operations with multiple information visualization techniques.

The initial surgical dashboard provided a strategic view of the department. An extension of the data warehousing layer is an observation engine which, when user specified events occur, will trigger processes ranging from communication (telecom, email, etc) to information exchange with other applications via web services. The user interface contains three major components: 1) A data manipulation layer which allows interaction with the data warehouse and provides analysts with means to create and track new metrics; 2) A visualization toolkit to create graphs and web pages to display information effectively. This includes the means to create clickable and animated graphs; and 3) A simplified means to specify observers within the business intelligence engine and canned solutions to communication information when events occur.

Objective 1.
Complete a business intelligence engine to handle aggregation and manipulation of data. Future work will entail the refinement of dashboards currently in beta testing.

To accommodate the need for data validation as well as tactical information about OR use from day to day an additional dashboard was developed. This dashboard was designed to focus on case to case problems of the perioperative environment. A greater scrutiny of daily performance and case data will drive the questions asked of the strategic dashboard which remains in beta testing. During the year of this report, the following work was completed. A tactical dashboard was developed in conjunction with the strategic dashboard and has been deployed to an internally hosted server. The tactical dashboard is being used within the perioperative environment to evaluate OR utilization, scheduling workflow, and case data accuracy. Data are being validated at the case level and new processes for data entry are being designed to ensure the accuracy of the metrics within the strategic dashboard. A semantic relationship query mechanism to facilitate ubiquitous, dynamic filtering of data was developed for the strategic dashboard. The strategic dashboard's user interface has been updated to include an initial design for filtering of information as well as a means to create dashboards. The perioperative data are being validated through the use of the use of the tactical dashboard. This validation is necessary before the release of the strategic dashboard as poor data quality causes inaccuracies in aggregated statistics.

Informatics subgroup 2. Ergonomics/Human Factors

Additionally, within our established informatics research, for over two years we have continually identified the emergence of *Ergonomics/Human Factors(E/HF)* as a major sub-component interest within the existing aims of our contract. As research progressed with the overall Innovations in the Surgical Environment program and in particular with the Informatics pillar, attention was drawn to the importance of these factors to patient safety and effective training in surgical procedures. This emergence of these areas of interest and subsequent investigation of ergonomics/human factors has been consistently and formally reported in two of our annual conferences and most recently in quarterly and annual reports to USAMRMC. Several manuscripts related to our research in ergonomics and human factors are placed in the Appendix to this report.

Human factors and Ergonomics are two related branches of study that examine the relationship between people and their work environment. Ergonomics often focuses on the physical environment and the human body, while human factors center more on the cognitive aspects of performance—how an operator interacts with the information environment. The same ergonomics and human factors techniques credited with making industrial processes safer and more efficient can be applied to the analysis and improvement of OR operations.

Our Informatics research pillar comprises and subsumes the investigation of ergonomics and human factors. To this point, our discussion of workflow has taken a macro or panoramic view; for example, how might we most effectively track and bring together the people and assets necessary to ensure that a patient's surgical experience is safe and efficient. Through our formal recognition of human factors and ergonomics within our existing research pillars, we focus on a more micro-level analysis, such as how the physical interface between the surgeon and the patient could be improved and the associated work space chaos and stressors of minimally invasive surgery (MIS) be reduced.

The patient is the center of the ORF. During MIS, the interfaces between the patient and the surgeon are critical to both the safety and quality of patient care and surgeon welfare. Patient-surgeon interfaces are complicated by compromises in equipment design, technology limitations, operating theatre layout, and technical approaches. In particular, ergonomic problems in the MIS workspace, such as obstructing catheters and cluttering tubes, can elevate the chance for contamination, increase surgical risks to the patient, and reduce work efficiency. Optimal workflow during MIS stands to be achieved through better understanding of patient-surgeon interfaces, both intracorporeal and extracorporeal. In the ORF, advanced technology could function as a key enabler, allowing an optimal patient-surgeon interface.

Some of our current work is focused on establishing quantitative, valid measures of workflow within patient-surgeon interfaces, identifying ergonomic problems that result as a consequence of workplace designs (e.g., arrangement or management of cables and catheters), and demonstrating key barriers to optimal workflow that present direct safety and efficiency concerns. One project is based on collaboration between surgical experts

and human factors experts. Previous experiences in video capturing and analysis are being used as a basis for development of workflow measures and identification of ergonomic inadequacies. Time-motion studies have been conducted to collect objective data on activities in the patient-surgeon interface. Conceptual workplace layout designs are being developed based on objective data and simulations of what workflow might be if interfaces were optimized.

Given the physical risks associated with performing laparoscopic surgery, ergonomics to date has focused on the primary minimally invasive surgeon. Similar studies have not extended to other operating room staff. Simulation of the assistant's role as camera holder and retractor during a Nissen fundoplication allowed investigation of the ergonomic risks involved in these tasks. Specific tasks to be completed in support of this research were identified as objectives of the study.

Objective 1. Continue to develop an assessment of difficulty hierarchy of Fundamentals of laparoscopic Surgery (FLS) tasks. This task will require extended work.

Objective 2. Develop an assessment of the effectiveness of self-mentored surgical training. We realized the importance of the fundamental understanding about the characteristics of the movement patterns utilized by expert laparoscopic surgeons. For the early stage of this particular research project, we have started establishing quantitative and objective methodologies to identify these expert movement patterns which must be substantially different than the movement patterns used by less experienced surgeons. We are also in the process of defining finite numbers of sub-movements which may create complex surgical movements by successive combination of several sub-movements.

Informatics subgroup 3. Context Aware Surgical Training (CAST)

We proposed to design and implement a prototype context aware surgical training environment (CAST) as part of the University of Maryland Medical System's SimCenter. This system was designed to explore the role that an intelligent pervasive computing environment can play to enhance the training of surgical students, residents and specialists. The research built upon prior work on context aware "smart spaces" done at UMBC; leverage our experience in working with RFID in the DARPA Trauma Pod program as well as in incorporating Web-based infrastructure and software applications in academic and professional development programs. The project was expected to result in a pilot system integrating one or two training resources available in the SimCenter into a context aware training environment that can recognize the presence of a trainee and or mentor and take appropriate action based on known training goals and parameters. The project proposed to advance the knowledge of context aware training environments in a highly technical medical field and provide a basis for incorporating more advanced technology assisted learning experiences in medicine. This "smart environment" may then, if successful, be scaled to meet the needs of an operative environment where the technological demands may be the similar or analogous to those seen in the training environment. Ultimately, the advanced training and potential for use in perioperative environments have a long-term end goal of improving patient safety and adding to the

body of knowledge in surgical training. Initially, we saw a situation were clinicians in training can receive a tailored curriculum. Additionally, we envisioned a system that offers real-time feedback and decision support and education metrics to faculty.

A key goal this year was to prototype the CAST system, and we defined a typical use case for our system. A Student enters the simulation center. The system identifies the student (for instance, using their Bluetooth phone or their badge), and does a prerequisite check based on the simulator the student wants to perform the procedure. Only if the student is done with the prerequisites, is he/she allowed to proceed. When the student indicates that they are ready to begin, the system starts capturing the external and internal view until the student indicates that they have completed the task. The captured video is then transferred to the video server for review by the instructor. The instructor interface allows the instructor to see the entry logs of students in terms of when they entered and exited the center along with the corresponding external view.

We employed the spiral prototyping approach as an experimental test bed; we designed and implemented an initial system prototype that would meet the above functional requirements. The prototype integrates two machines with each simulator -- a small Nokia 800 device for resident interaction, and a larger PC for video capture. Note that this is for the proof of concept. A single small form factor but computationally powerful machine could be used instead. In fact, for virtual reality (VR) simulators we expect that manufacturers could eventually integrate our system directly into the computer that drives the simulation.

Our prototype used Bluetooth for localization of residents in the simulation center. It was designed to be modular, so that any other technology (such as resident ID cards) could be integrated easily. We also hosted training materials including videos for FLS, Kentucky and Rosser tasks in our system, and tracked student progress through the chapters checked out. This was used for enforcing prerequisites when students entered the simulation centre to perform procedures. In addition to enforcing prerequisites, there was a need for the instructors to visually see what the residents were doing during their simulation procedures. We use N800's built in camera to capture the residents' external views. These video feeds are then fed into a central server for review by the instructor.

For location detection, we also experimented with using the Awarepoint tags. Awarepoint uses a zigbee based mesh network for localization and exposes the location information through a web service. Our experiments indicated that Awarepoint could provide us room level information, but not anything finer. While this would help identify if the residents were in the simulation center, it would not help determine which machine they were using, which was needed for CAST. We demonstrated our first system prototype at the ORF workshop by going through a typical student workflow.

We also focused on moving the system from UMBC machines to the MASTRI infrastructure where they will be housed. We purchased a small factor Dell machine to be used for capturing internal views from simulators. Storage was purchased and added to the mastri-internal server for archiving both internal and external video feeds. Also, we

have integrated the student database from the hospital, hosted FLS and other training videos on the hospital infrastructure and hacked internal views of the simulators. We developed the system to capture internal video feeds and metrics from the following simulators; Promis, Stryker and the Laproscopic VR simulator.

Current efforts have focused on testing an initial deployment of the CAST system at the MASTRI Center. We demonstrated the system to a set of resident volunteers for feedback in a form of Beta-test of the system. We set up hardware and software to include the VR Simulator as part of the CAST system deployment. We got usability feedback and fixed bugs. A significant part of the effort was also spent in surveying the state-of-the-art in Video/VR usage for surgical training. We identified a small but significant body of work (e.g. Sinanan et al, Darzi et al) in checking the construct validity of the models for training using these simulation tools. The typical approach is to use sensors to capture the kinematics of the tools, as well as force/torque measures. The UMBC/MASTRI team decided that we would like to focus on an alternate approach that i) focused on the video, not (initially) any other sensors and ii) tried to capture using machine learning techniques the ability of an expert surgeon to identify key events in a surgery that relate to outcome or skill assessment. This is a very challenging and open problem. Key initial steps were identified for initial implementation in the first year.

A detailed description of the CAST project, "A Ubiquitous Context-Aware Environment for Surgical Training", was presented at the First International Workshop on Mobile and Ubiquitous Context Aware Systems and Applications (MUBICA 2007), August 2007, by P. Ordóñez, P. Kodeswaran, V. Korolev, W. Li, O. Walavalkar, B. Elgamil, A. Joshi, T. Finin, Y. Yesha, I.George.   This presentation is contained in the Appendix to this report.

NOTE:

The effort to establish a system for archiving imaged data from training sites has been attenuated due to advancements in archiving capability in off-the-shelf systems. The focus of this project now rightly shifts to video summarization by unique application of artificial intelligence techniques. Video summarization has extraordinary potential for streamlining the events in the future perioperative environment. Further, there are many and varied military applications from video summarization. The UMBC Graduate students currently working on the CAST and the background research effort for the new direction will transition out, as the new direction is less closely aligned with their research interests. Dr. Mike Grasso, MD/PhD in Computer Science, will be joining the effort, and a new graduate student whose research focus will be on the video efforts will join the team. This of course is subject to new funds being available.

Informatics subgroup 3. Video Summarization

Our overall goal is to identity key portions of surgical procedures to aid in video-based assessment. To establish feasibility, we set out to identify the critical view of a laparoscopic cholecystectomy. The critical view is used to identify the key anatomy after major dissecting has been completed, but before clipping the cystic duct and artery.

During the past year, we completed an initial analysis of this problem. We compared more than 50 image features with a distance metric to identify the critical view of a laparoscopic cholecystectomy. We experimented with roughly 50 different images features and several distance metrics. Our initial results showed a 72% sensitivity and 72% specificity. The study was small in size, using only 5 laparoscopic cases, and our comparisons were limited to one image feature at a time. An abstract was submitted to the American Medical Informatics Association (AMIA) Fall Symposium (appendicized).

During the last several months, we have been working two new initiatives. The first is the creation of an image classifier using a support vector machine (SVM). This is a machine learning approach that uses multiple image features to train the image classifier. Our initial accuracy with the SVM improved to about 90%. This original SVM was built from only 5 laparoscopic cases. Our understanding is that 25 additional cases will be available after IRB approval has been obtained. Accuracy should improve when more cases are used to train the SVM. In addition, we used particle analysis and edge detection to identify key segments inside each image. We also plan to use this data to increase the accuracy of the SVM. During this quarter, we prepared a protocol for the use of 25 additional video cases upon which to build the SVM.

Informatics subgroup 4. Operating Room Clutter (ORC)

The Operating Room Clutter project enters its final phase under the provisions of the contract, and ended during the period of performance reported here. Further research activities will seek support from other funding agencies.

Prior to completion, the project team worked on the use of advanced video technology to support coordination in operating rooms. Activities were in four areas. All publications referred to may be found in the website: http://hfrp.umaryland.edu. For full length journal articles, PDF files may be downloaded. For others, abstracts are available. In all, we published 8 full-length peer reviewed journal articles, 2 full-length peer reviewed proceeding articles, and 8 conference abstracts. The references below can provide further details.

> **A. Models of decision making for operating room management**.
> We reviewed literature and developed a synthesis report on the state of the art of decisions on the day of surgery. Furthermore, we developed models for decision support systems for operating room management. The activities in this area were reported in the following publication:
>
> 1. Dexter F, Xiao Y, Dow AJ, Strader MM, Ho D, Wachtel RE. Coordination of Appointments for Anesthesia Care Outside of Operating Rooms Using an Enterprise Wide Scheduling System. *Anesthesia and Analgesia. 105:1701-1710. 2007*

B. Operating room multimedia system design and methodology.
We developed technology, primarily based on algorithms of video processing and biosignal processing, to display status of operating rooms. The displays are to increase situational awareness. The technological advances made by our group were reported in the following publications:

2. Xiao Y, Schimpff S, Mackenzie CF, Merrell R, Entin E, Voigt R, Jarrell B. Video Technology to Advance Safety in the Operating Room and Perioperative Environment. *Surgical Innovation. 14(1): 52-61. 2007*

3. Hu P, Xiao Y, Ho D, Mackenzie CF, Hu H, Voigt R, Martz D. Advanced Visualization Platform for Surgical Operating Room Coordination: Distributed Video Board System. *Surgical Innovation. 13(2):129-135. 2006*

4. Hu P, Seagull FJ, Mackenzie CF, Seebode S, Brooks T, XiaoY. Techniques for Ensuring Privacy in Real-Time and Retrospective Use of Video. *Telemedicine and e-Health, 12(2): 204, T1E1. 2006*

C. Survey and descriptive studies of operating room management, with and without the support of advanced video technology.
In conjunction with technology development, we conducted observational and survey studies of operating room management. These studies and associated results were in the following publications:

5. Seagull FJ, Xiao Y, & Plasters C. Information Accuracy and Sampling Effort: A Field Study of Surgical Scheduling Coordination. *IEEE Transactions on Systems, Man, and Cybernetics, Part A:Systems and Humans. 24(6), 764-771. 2004*

6. Dutton R, Hu PF, Mackenzie CF, Seebode S, Xiao Y. A Continuous Video Buffering System for Recording Unscheduled Medical Procedures. *Anesthesiology, 103:A1241. 2005*

7. Gilbert TB, Hu PF, Martz DG, Jacobs J, Xiao Y. Utilization of Status Monitoring Video for OR Management. *Anesthesiology, 103:A1263. 2005*

8. Dutton R, Hu P, Seagull FJ, Scalea T, Xiao Y, . Video for Operating Room Coordination: Will the Staff Accept It?. *Anesthesiology: 101: A1389. 2004*

D. Technology evaluation.
We conducted evaluation studies of the technology deployed. The primary focus was on user acceptance and usage patterns. The focus was chosen because the current science of operating room management has concluded that improvement of decision making on the day of surgery will lead to improvement in intangible outcomes, such as situation awareness, and will unlikely lead to improvement in

operating room throughput (e.g., volumes and economic returns). Our work was reported in the following publication.

9. Xiao Y, Dexter F, Hu FP, Dutton R. <u>Usage of Distributed Displays of Operating Room Video when Real-Time Occupancy Status was Available</u> . Anesthesia and Analgesia 2008; 106(2):554-560. 2008

10. Kim Y-J, Xiao Y, Hu P, Dutton RP. <u>Staff Acceptance of Video Monitoring for Coordination: A Video System to Support Perioperative Situation Awareness</u>. *Journal of Clinical Nursing (accepted). 2007*

The project team has worked on the use of advanced video technology to support coordination in operating rooms. We developed models for decision support systems for operating room management. We developed technology, primarily based on algorithms of video processing and biosignal processing, to display status of operating rooms.

In conjunction with technology development, we conducted observational and survey studies of operating room management. We conducted evaluation studies of the technology deployed. The primary focus was on user acceptance and usage patterns. The focus was chosen because the current science of operating room management has concluded that improvement of decision making on the day of surgery will lead to improvement in intangible outcomes, such as situation awareness, and will unlikely lead to improvement in operating room throughput (e.g., volumes and economic returns).

Informatics subgroup 5. Improving Perioperative Communications (IPC)

Background:

In the UMMS OR the Cardiac Surgery Service utilizes a common communications point (a "cardiac phone line") that in a sense is used to acquire information and provide that information to any team member who calls the line to acquire information. The cardiac phone line has been scripted and is actively in use through a voice mail system. It can only be altered by dedicated personnel with password capability. The script involves the following standardized information: Identification of individual providing information, the Date of surgery, the Total number of cases, and OR location, patient name, case order, medical record number, age, surgeon, anesthesiologist and procedure. Evening schedule updates have been made possible through a second phone line option.

After some effort, we can now move to track updates on the phone line and correlate these updates with OR start delays. Thus, we refined the IPC question to Does more accurate information as evidenced by updates on the phone line, ie improved communication, result in fewer problems in the morning with cardiac surgical cases starting on time- are instruments better prepared for the procedure, are operating rooms better equipped for the appropriate case, are the correct pick lists utilized for the correct surgeon, is there less of a transport delay because the patient's hospital location has been

identified? The question contains reference to some of the delay codes that are currently utilized by the Operating room tracking system and reported for glitch analysis.

With the assistance of the communications personnel, we reconfigured the cardiac phone line so that we can actually track the phone calls made to the phone line. This enabled us to: Determine key personnel who are utilizing the phone line, Determine groups of personnel utilizing the phone line (i.e. nursing, anesthesia, perfusion), Determine which groups are not utilizing phone line information (i.e. anesthesia techs), Determine whether there is a time variable; is there a better time to call for updates? Should updates be made at predetermined times or should they be more dynamic?

We hypothesized that information gained from increased communication improves OR efficiency. If this is the case we can then move to see if more real-time enabling technologies might be deployed to other services within the UM ORs and perhaps other ORs "everywhere".

During this period of performance, the team sought to find an appropriate "question" with which to focus this effort. In particular, there was a need to tie a performance metric (perioperative workflow related) to the IPC task. Although some progress was made, this aspect of Innovations in the Surgical Environment will end and future efforts will be refocused. Funding will be sought from other agencies

B. Simulation

Simulation.1 The Maryland Virtual Patient

We present here a simplified description of the MVP simulation, interaction and tutoring system. A virtual patient instance is launched and starts its simulated life, with one or more diseases progressing. When the virtual patient develops a certain level of symptoms, it presents to the attending physician, the system's user. The user can carry out, in an order of his or her choice, a range of actions: interview the patient, order diagnostic tests, order treatments, and schedule the patient for follow-up visits. The patient can also automatically initiate follow-up visits if its symptoms reach a certain level before a scheduled follow-up. This patient-physician interaction can continue as long as the patient "lives."

As of the time of writing, the implemented MVP system includes a realization of all of the above functionalities, though a number of means of realization are temporary placeholders for more sophisticated solutions, currently under development. The most obvious of the temporary solutions is the use of menu-based patient-user interaction instead of natural language interaction. While this compromise is somewhat unnatural for our group, which has spent the past 20 years working on knowledge-based NLP, it has proved useful in permitting us to focus attention on the non-trivial core modeling and simulation issues that form the backbone of the MVP system.

MVP currently covers six esophageal diseases pertinent to clinical medicine: achalasia, gastroesophageal reflux disease (GERD), laryngopharyngeal extraesophageal reflux disease (LERD), LERD-GERD (a combination of LERD and GERD), scleroderma esophagus and Zenker's diverticulum.

At the beginning of a simulation session, the system presents the user with a virtual patient about whose diagnosis he initially has no knowledge. The user then attempts to manage the patient by conducting office interviews, ordering diagnostic tests and prescribing treatments.

Answers to user questions and results of tests are stored in the user's copy of the patient profile, represented as a patient chart. At the beginning of the session, the chart is empty and the user's cognitive model of the patient is generic – it is just a model of the generalized human. The process of diagnosis results in a gradual modification of the user's copy of the patient's profile so that in the case of successful diagnosis, it closely resembles the actual physiological model of the patient, at least, with respect to the properties relevant to the patient's complaint. A good analog to this process of gradual uncovering of the user profile is the game of Battleship, where the players gradually determine the positions of their opponent's ships on a grid.

At any point during the management of the patient, the user may prescribe treatments. In other words, the system allows the user not only to issue queries but also to intervene in the simulation, changing property values within the patient. Any single change can induce other changes – that is, the operation of an agent can at any time activate the operation of another agent.

Simulation: Utility

The MVP project can be viewed as just one of a number of applications in the area of intelligent clinical systems. The latter, in turn, can be viewed as one of the possible domains in which one can apply modeling teams of intelligent agents featuring a combination of physical system simulation and cognitive processing.

So, in the most general terms, our work can be viewed as devoted to creating working models of societies of artificial intelligent agents that share a simulated "world" of an application domain with humans in order to jointly perform cognitive tasks that have until now been performed exclusively by humans. Sample applications of such models include:

- o a team of medical professionals diagnosing and treating a patient (with humans playing the role of either a physician or a patient)
- o a team of intelligence or business analysts collecting information, reasoning about it and generating analyses or recommendations (with humans playing the role of team leader)

o a team of engineers designing or operating a physical plant (with humans playing the role of team leader)
o a learning environment (where humans play the role of students).

As can be seen, this work is at the confluence of several lines of research – cognitive modeling, ontological engineering, reasoning systems, multi-agent systems, simulation and natural language processing.

During the period of performance, we have been working on the following issues:

1. We have continued to develop a computational model of the cognitive agent. We have tested the goal- and plan-based reasoning component and its interaction with the interoceptive and language perception modules and verbal, mental and physical action simulation modules.

2. We have spent much of the time preparing for the demonstration of the system at the program conference and at several meetings of the American College of Surgeons. In particular, we have developed a new demo interface.

3. We have continued to work on the natural language substrate of the system, concentrating on enhancements required for processing dialog (not expository text). As part of this module, we have implemented an enhanced microtheory of indirect speech acts.

4. We have continued working reference resolution algorithms (this is a very difficult task in and of itself).

5. We have continued work on the acquisition of ontology and lexicon knowledge.

6. Improvement of the DEKADE user interface has continued apace. New facilities for editing and viewing intermediate and final results of text analysis have been introduced and existing ones improved.

7. We have spent considerable time on improving the documentation of the project work. We have written and submitted for publication 5 papers describing aspects of our system.

During this reporting period, the research team continued to refine the cognitive simulation system by adding more clinical scenarios and challenges. The communication between patient and physician has reached a high level of realism and clinical utility.

An extensive summary of this project's work was presented at the annual meeting of the American College of Surgeons, at the DOD-sponsored workshop on psychometrics of simulation/games, and a similar DOD meeting in San Antonio, Texas.

We continue to discuss with representatives of the DOD Medical Departments the application of cognitive simulation training for far-forward care providers. As this project winds down, the investigators will seek sources of funding to attain the potential of the cognitive simulation system.

Simulation.2 Training for Surgical Excellence and Patient Safety

The development of the Maryland Advanced Simulation, Training, Research and Innovation center (MASTRI) opened the door to innovative research opportunities that enhance surgical training and improve patient safety. Within the existing scope of the current contract, several projects will be undertaken during the final year of the contract that conceive, develop and validate simulation-based training for proficiency in the performance of surgical tasks.

For the milestone pertaining to the exploration of the application of technologies to refine methods of medical instruction, the following activities took place:
- Online training system: we have developed offline resources for training laparoscopic cholesystectomy procedures involving the use of video vignettes. These resources are slated for implementation using an online learning system. We have identified a doctoral candidate in computer science to facilitate implementation of the system.
- Audience polling system: We have recently acquired a system for audience-response measurement. Our system for audience-response measurement is now integrated into the training of all residents for polling response to training and medical grand rounds. This system will also be used to facilitate and enhance our presentation at the annual meeting of the Society for Simulation in HealthCare.
- Simulation-based competency based training: We are currently refining and expanding our use of criteria-based training, which uses measures of performance to determine training sufficiency. We are currently refining the criteria for Virtual reality (VR) and physical-model training for basic laparoscopic skills.

For the milestone related to developing new models for simulation training:
- Arthroscopic simulation: Arthroscopic skills models continue to be refined. A new model for spinal disk herniation has been developed in the form of a prototype. Pilot testing of the model's functions has been carried out. Curriculum for the model is being developed.
- Ventral hernia: Provisional patent obtained. Prosecution of full patent is in process. Pilot validation trials of our model have been carried out by another university. The simulator is now in routine use as a teaching tool for fellows, residents, and industry representatives, and provides a cost-effective alternative to porcine model of ventral hernia repair. Validation trials are being designed.
- Suture-skills drill: A new physical model of tissue and pattern of visual targets marked upon the tissue was developed for a flexible curriculum of training suturing skills. We are in the process of refining models for developing the curriculum for drills to practice suturing skills.

For the milestone related to configuring the training site for OR, ICU, Emergency Department and Team Training scenarios:

New equipment and infrastructure were acquired during this reporting period. Specifically, additional simulation equipment to be used for a variety of procedural and skill-based instruction has been delivered. The Sim baby simulator has been obtained, a pediatric exam bed has been integrated into the simulation space, ultrasound equipment was obtained, allowing ultrasound-guided catheter placement training, curtains and storage have been installed, and new part-task trainers have been obtained.

- Significant progress has been made in the configuration of and use of OR B as a training site. Video and audio cabling and networking hardware has been installed and all communication systems are near completion.
- Hiring of a simulation educator/technician has brought us to near full operation of the OR B training site.
- Many new full and part-task trainers have allowed the beginning of several courses with a large number waiting in the wings. More than a dozen new courses for medical and surgical residents, nursing personnel and medical students have been introduced using this new training site.

For the development of collaborative ventures with academic institutions and government agencies, we are currently preparing extramural grant proposals on the topic of developing metrics for surgical training, specifically regarding the assessment of medical resident performance. Substantive collaboration is dependent on receipt of funding.

C. Smart Image

C.1 Smart Image: CT-guided Imaging

Having completed the entirety of the experimental and animal imaging work, our efforts remained focused on data processing and scientific reporting during the current year. The following were significant notable outcomes.

1. We made an oral presentation on our work on Live AR at the SAGES conference help in Phoenix, AZ in April 2009. Subsequently, an in-depth manuscript on the same topic was submitted to Surgical Endoscopy, the official journal of SAGES, for possible publication. **The submitted manuscript is placed in the Appendix.** This reporting period included work on refining and producing results, especially Live AR movie clips, for the conference presentation and the manuscript.

2. We also presented a poster at the Computer-Assisted Radiology and Surgery (CARS) conference in Berlin, Germany in June 2009. This presentation explored the specific topic of using image registration for continuous volumetric CT-guided interventions. **A copy of the abstract is presented in the Appendix.** A manuscript will follow this presentation.

3. We continued to work on low-dose CT reconstruction, one of the originally proposed technical objectives. We worked with Philips, the manufacturer of our CT scanner, on data preprocessing issues that now enable us to reconstruct images with a consistent orientation and thereby allow head-to-head comparison of reconstructed images when x-ray dose is varied. A manuscript summarizing the work will be prepared and submitted in the upcoming year.

C.2. Smart Image: Image Pipeline

The work we have completed in the past year is summarized in a number of peer-reviewed publications and has been shared in several presentations. These are discussed in the context of our overall goals – to develop visualization requirements, principles, and frameworks, as well as solutions to specific computational challenges, which will permit useful and usable augmentation of the laparoscopic image. This augmentation assumes input from other modalities, including surgeons' annotations, as well as pre-operative and intra-operative CT and other imaging techniques. The full papers are in the appendix.

Visualization Framework:

We have described the overall visualization framework for the Smart Image project, including both the computational and usability components, in a paper submitted to *Surgical Innovations*:

1) Yang, R., Carswell, C.M., Wang, X., Zhang, Q., Han, Q., Lio, C., and Seales, B. Mapping the Way to a Dual Display Framework for Laparoscopic Surgery.

Abstract:

Many performance and workload problems associated with the use of traditional laparoscopic displays are the result of spatial disorientation. This premise has guided our development of a dual display framework for computer-augmented surgical displays, allowing us to take guidance from research on how to design successful navigation aids (navaids) for large-scale environments.   Our dual-display combines the traditional scope (forward track) view with a computationally-generated global 3D (map) view. The latter provides a wider field of view, explicit cues to depth and scale, and a way to view interior and exterior surfaces of target anatomy from different approach angles. One way to implement such a 3D view is to extract images of surface textures from a laparoscopy video sequence and then map the texture onto pre-built 3D objects, for example surface models derived from MR/CT. We describe an algorithm that takes advantage of the fact that nearby frames within a video sequence usually contain enough coherence to allow 2D-2D registration, a much better understood problem than 2D-3D registration. Our texturing process can be bootstrapped by an initial 2D-3D manual-assisted registration of the first video frame followed by mostly-automatic texturing of subsequence frames. Initial research on the validity of our technical approach indicates that it improves

registration performance compared to a standard registration technique that relies on camera tracking. Ongoing technical and usability evaluations of the system are being conducted in order to ensure system functionality.

Front end assessments for requirements and acceptability are described in a peer-reviewed proceedings paper presented at the 2009 Human Factors and Ergonomics Society meeting.

2) Lio, C.H., Carswell, C.M., Han, Q., Park, A., Strup, S., Selaes, W.B., Clarke, D., Lee, G., and Hoskins, J. (2009). Using Formal Qualitative Methods to Guide Early Development
of an Augmented Reality Display System for Surgery. Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting. Santa Monica, CA: HFES.
Abstract:

Nine laparoscopic surgical experts (2 residents, 4 fellows, and 3 surgeons) underwent semi-structured interview questions to evaluate the concept of a "dual-view" display for laparoscopic surgery. The 30-40 minute audio-recorded interviews were transcribed, submitted to an open source qualitative program for classification and categorizing, and were condensed for the iterative processes of analysis and interpretation. Findings revealed that despite the relatively brief interview sessions and limited number of surgical experts available, the experts provided sufficient insights and suggestions to guide further development of prototypes. This means that the use of semi-structured interviews as an expert knowledge elicitation technique may be suitable for assessing the development of augmented reality display systems for surgical and training applications, and it may have promise for the development of augmented and virtual environments more genially.

Computational Challenges:

The past year also saw the publication of a peer-reviewed journal article summarizing the procedure we have developed for registering a series of video images from the laparoscope to prebuilt surface models without using camera tracking.

3) Want, X., Zhang, Q., Han, Q., Yang, R., Carswell, M., Seales, B., and Sutton, E. (2009) Endoscopic video texture mapping on pre-built 3D anatomical objects without camera tracking. IEEE Transactions on Medication Imaging, 7(7), 1-12.

Abstract:

Traditional minimally invasive surgeries use a view port provided by an endoscope or laparoscope. We argue that a useful addition to typical endoscopic imagery would be a global 3D view providing a wider field of view with explicit depth information for both the exterior and interior of target anatomy. One technical challenge of implementing such a view is

finding efficient and accurate means of registering texture images from the laparoscope on pre-built 3D surface models of target anatomy derived from magnetic resonance (MR) or computed
tomography (CT) images. This paper presents a novel method for addressing this challenge that differs from previous approaches, which depend on tracking the position of the laparoscope. We
take advantage of the fact that neighboring frames within a video sequence usually contain enough coherence to allow a 2D-2D registration, which is a much more tractable problem.
The texturing process can be bootstrapped by an initial 2D-3D user-assisted registration of the first video frame followed by mostly-automatic texturing of subsequent frames. We perform
experiments on phantom and real data, validate the algorithm against the ground truth, and compare it with the traditional tracking method by simulations. Experiments show that our
method improves registration performance compared to the traditional tracking approach.

We also published a peer-reviewed proceedings paper on a method for acquiring and reconstructing 3D surface models based on light fall-off between the camera and organ surface.

4) Liao, M., Wang, L., Yang, R., and Gong, M. Real-time light fall-off stereo. (2008). International Conference on Image Processing (ICIP).

Abstract:

We present a real-time depth recovery system using Light Fall-off Stereo (LFS). Our system contains two co-axial point light sources (LEDs) synchronized with a video camera. The
video camera captures the scene under these two LEDs in complementary states(e.g., one on, one off). Based on the inverse square law for light intensity, the depth can be directly solved using the pixel ratio from two consecutive frames. We demonstrate the effectiveness of our approach with a number of real world scenes. Quantitative evaluation shows that our
system compares favorably to other commercial real-time 3D range sensors, particularly in textured areas. We believe our system offers a low-cost high-resolution alternative for depth
sensing under controlled lighting.

Yet another peer-reviewed proceedings paper was recently presented at MICAI (Medical Image Computing & Computer Assisted Intervention) and dealt with the methods to model the intra-object deformations with using a small number of parameters that can be applied to new target objects. This allows for better registration of pre-built 3D shape models of target organs to their corresponding laparoscopic video sequence.

5) Han, Q. Strup, S., Carswell, C.M., Clarke, D., Seales, W.B. (2009). Model Completion via Deformation Cloning Based on an Explicit Global Deformation Model. 12th International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)(1) 2009: 1067-1074.

Abstract:

Our main focus is the registration and visualization of a pre-built 3D model from preoperative images to the camera view of a minimally invasive surgery (MIS). Accurate estimation of soft-tissue deformations is key to the success of such a registration. This paper proposes an explicit statistical model to represent global non-rigid deformations. The deformation model built from a reference object is cloned to a target object to guide the registration of the pre-built model, which completes the deformed target object when only a part of the object is naturally visible in the camera view. The registered target model is then used to estimate deformations of its substructures. Our method requires a small number of landmarks to be reconstructed from the camera view. The registration is driven by a small set of parameters, making it suitable for real-time visualization.

Continuing Work:

We plan to have shipped the completed dual display simulation (interactive prototype for usability testing) to UMMC by November 15. This allows for the precise assessment of the presumed advantages of our dual display framework for navigation and the development of adequate spatial situation models. It is our hope that the framework will form the basis for continued collaboration and development of user-centered visualization systems for minimally invasive surgery.

Key Research Accomplishments

A. Informatics

Informatics subgroup 1. Perioperative Scheduling Study

Major Accomplishments achieved during this period of performance include the development of a mathematical congestion evaluation model for evaluating congestion in post-operative units, including ICUs, IMCs, and floor units. This model requires data about post-operative destinations and length-of-stay distributions for different types of surgeries. We analyzed data about cardiac surgeries from two years and have analyzed UMMC financial records for all of the surgical cases for a year.

Informatics subgroup 2. Operating Room Glitch Analysis

A tactical dashboard was developed in conjunction with the strategic dashboard and has been deployed to an internally hosted server. The tactical dashboard is being used within the perioperative environment to evaluate OR utilization, scheduling workflow, and case data accuracy. Data are being validated at the case level and new processes for data entry are being designed to ensure the accuracy of the metrics within the strategic dashboard.

Informatics subgroup 3. Context Aware Surgical Training (CAST)

A prototype CAST system was emplaced in the MASTRI system for assessment. Work was done to design a system of evaluation of the system in terms of improvements in learning outcomes due to self-feedback, improvements in learning outcomes due to instructor feedback and synchronous versus asynchronous feedback. We demonstrated the system to a set of resident volunteers for feedback in a form of Beta-test of the system. We set up hardware and software to include the VR Simulator as part of the CAST system deployment. We got usability feedback and fixed bugs.

Informatics subgroup 3. Video Summarization

We completed an initial analysis of this problem. We compared more than 50 image features with a distance metric to identify the critical view of a laparoscopic cholecystectomy. We experimented with roughly 50 different images features and several distance metrics. Our initial results showed a 72% sensitivity and 72% specificity. An abstract was submitted to the American Medical Informatics Association (AMIA) Fall Symposium.

Informatics subgroup 4. Operating Room Clutter (ORC)

During this period of performance, we published 8 full-length peer reviewed journal articles, 2 full-length peer reviewed proceeding articles, and 8 conference abstracts.

B. Simulation (Virtual Patient)

n this final year of the project, our team has delivered additional versions of the Maryland Virtual Patient Environment. The realism of the simulation has been enhanced by including coverage of "unexpected" interventions; allowing discontinued treatments; allowing new diseases to develop due to side effects of treatments. The user interface has been redesigned. A new agent-based architecture has been developed to support enhanced cognitive capabilities of the virtual patient and the intelligent tutor, including language capabilities. In the area of language processing, a dialog processing model was developed. Work has continued on improving the language understanding capabilities, centrally including treatment of referring expressions. Enhancement of static knowledge resources, the ontology and the lexicon, has been ongoing. Work on extending the coverage of diseases has been ongoing: a further improvement of the model of GERD is under way, as is the modeling of cardiovascular diseases. A totally reworked system version, with dialog support, was released in June 2008. Work has also been ongoing on improving and extending the set of development tools – the DEKADE demonstration, evaluation and knowledge acquisition environment supporting natural language work has been revamped; the interface for creating instances of virtual patients has also been enhanced; a web-based environment for supporting internal documentation has been installed.

B. Simulation (Training for Surgical Excellence)

For the milestone related to configuring the training site for OR, ICU, Emergency Department and Team Training scenarios, new equipment and infrastructure were acquired during this reporting period.   Specifically, additional simulation equipment to be used for a variety of procedural and skill-based instruction has been delivered. The Sim baby simulator has been obtained, a pediatric exam bed has been integrated into the simulation space, ultrasound equipment was obtained, allowing ultrasound-guided catheter placement training, curtains and storage have been installed, and new part-task trainers have been obtained.   Significant progress has been made in the configuration of and use of OR B as a training site. Video and audio cabling and networking hardware has been installed and all communication systems are near completion.   Hiring of a simulation educator/technician has brought us to near full operation of the OR B training site.   Many new full and part-task trainers have allowed the beginning of several courses with a large number waiting in the wings. More than a dozen new courses for medical and surgical residents, nursing personnel and medical students have been introduced using this new training site.

C. Smart Image

C.1. Smart Image: CT guided imaging

We made an oral presentation on our work on Live AR at the SAGES conference help in Phoenix, AZ in April 2009. Subsequently, an in-depth manuscript on the same topic was submitted to Surgical Endoscopy, the official journal of SAGES.   We also presented a poster at Computer-Assisted Radiology and Surgery (CARS) conference in Berlin, Germany in June 2009. This presentation explored the specific topic of using image registration for continuous volumetric CT-guided interventions. We continued to work on low-dose CT reconstruction, one of the originally proposed technical objectives. We worked with Philips, the manufacturer of our CT scanner, on data preprocessing issues that now enable us to reconstruct images with a consistent orientation and thereby allow head-to-head comparison of reconstructed images when x-ray dose is varied. A manuscript summarizing the work will be prepared and submitted in the upcoming year.

C.2. Smart Image: Image Pipeline

The work we have completed in the past year is summarized in a number of peer-reviewed publications and has been shared in several presentations. These are discussed in the context of our overall goals – to develop visualization requirements, principles, and frameworks, as well as solutions to specific computational challenges, which will permit useful and usable augmentation of the laparoscopic image. This augmentation assumes input from other modalities, including surgeons' annotations, as well as pre-operative and intra-operative CT and other imaging techniques.

Reportable Outcomes

We advanced the body of knowledge pertaining to informatics, smart image, simulation and human factors as these relate to surgical procedures, the perioperative environment

and the training of surgery.   We published more than forty manuscripts, hosted national and international meetings related to innovation in the surgical environment, and incorporated technical advances into patient care in a large academic medical center. We influenced significantly the training of more than three hundred fellows and residents, hundreds of staff and care providers and numerous medical students.

Perhaps our most important accomplishment has been the identification of a new of basic surgical sciences. These include computer and physical sciences, informatics, smart imaging, simulation and ergonomics and human factors that underpin surgical training. This event is a landmark of sorts, as it has changed forever the course of surgical education. Lessons learned from this research effort are being applied in training programs throughout the country and internationally.

Conclusion

This report began with the recognition that an extraordinary evolution in surgical care has occurred caused by rapid advances in technology and creative approaches to medicine. The increased speed and power of computer applications, the rise of visualization technologies related to imaging and image guidance, improvement in simulation-based technologies (tissue properties, tool-tissue interaction, graphics, haptics, etc) have interacted to advance the practice of surgery. However, the medical profession lags behind other applications of information systems. The research program reported here has proceeded under the mantle of "Operating Room of the Future". As a natural occurrence in the outcome of lessons learned in medicine, we are replacing that theme with the more appropriate "Innovations in the Surgical Environment."

This research program has consisted of three major pillars; OR informatics, simulation, and smart image. This year, we incorporated the research focus areas of physical and cognitive ergonomics and human factors into the informatics pillar. A summary description of the entire research portfolio was included in the appendix.
\
The purpose of the OR informatics program is to develop, test, and deploy technologies to collect real-time data about key tasks and process elements in clinical operating rooms. We have established testbeds of activities in both simulated and operational environments. We are currently performing tests of the hardware, refining software, and applying lessons learned to hospital operational functions. The objective of Simulation research is to create a system where a user can interact with a virtual human model in cognitive simulation and have the virtual human respond appropriately to user queries and interventions in clinical situations, with a focus on cognitive decision making and judgment. We have made significant strides toward realizing these goals. The MVP simulation functions well for esophageal disorders, and is continuing to expand the repertoire of diseases that are in the simulation model.

The objective of smart image is use real-time 3D ultrasonography and 40-slice highframe-rate computed tomography (CT) for intraoperative imaging to volume rendered anatomy from the perspective of the endoscope. We are combining CT and

Ultrasound to overlay image and data to enhance the performance of surgeons-in-training. We have carried out animate model testing of the image registration with great success. We continue to refine and expand our capability through hardware and software refinement.

In the future, OR workspace layout would be optimized through ergonomic data and human factors analysis, and this optimization would lead to the establishment of "best practices" for an array of surgical operations. Proper layout would reduce risks of infection, speed operations, and reduce fatigue of surgeons and staff, all elements that could contribute to a reduction in AEs and improved patient safety.

The year ahead is full of promise for refinements in the use of informatics to support safe and efficient operating room procedures, the use of simulation to improve and accelerate the training of competent surgeons, and the blending of imaging capabilities to provide clearer and safer interactions between patient and surgeon.

As stated earlier in this report, the current contract, W81XWH-06-2-0057, has been tied to a prior and topically related contract, DAMD-17-03-2-0001. The prior contract closed in February of 2009; the current one was scheduled to close in October of 2009; this contract was granted a no-cost extension until October, 2010. Some projects contained in the Informatics pillar, OGA, ORC and IPC, have been completed. The WORQ project will continue under the current contract as will the CAST project that has been reshaped into Video Summarization. Simulation for Training and the ergonomics/human factors work will continue through the period of no-cost extension.

These changes represent the maturing of a research endeavor over the course of six years, an endeavor which opened the door to a new set of basic surgical sciences. The Innovations in the Surgical Environment conference planned for the spring of 2010 will summarize the entirety of the research effort and point the direction to future innovative approaches to advance surgical technology in behalf of patient safety.

Publications

ABSTRACT: High-Speech Reconstruction of Low-Dose CT Using Iterative Techniques for Image-Guided Interventions. Doctoral Dissertation, University of Maryland 2008. Venkatesh Bantwal Bhat

ABSTRACT: High-Speech Reconstruction of Low-Dose CT Using GP-GPU. Venkatesh Bhat and Raj Shekhar

ABSTRACT: Live Augmented Reality for Laparoscopic Surgery Using a Novel Imaging Method – Initial Results from a Porcine Animal Model. R Shekhar PhD, CD Godinez, M.D., S Kavic, M.D., E Hart, M.D., Ivan George, AE Park, M.D.

Nagy P, Ramon Konewko, Max Warnock, Wendy Bernstein, Jacob Seagull, Yan Xiao, Ivan George, and Adrian Park. "Novel, Web-Based, Information-Exploration

Approach for Improving Operating Room Logistics and System Processes", Surg Innov 2008 15: 7-16.

Ordóñez P, P. Kodeswaran, V. Korolev, W. Li, O. Walavalkar, B. Elgamil, A. Joshi, T. Finin, Y. Yesha, I.George. "A Ubiquitous Context-Aware Environment for Surgical Training". The First International Workshop on Mobile and Ubiquitous Context Aware Systems and Applications (MUBICA 2007), August 2007.

Seagull FJ, Moses GR, Park AE. "Integration of Virtual Reality and Conventional Skills Trainers: A Mixed Resource Model." In Westwood J.D., Haluck R.S., Hoffman H.M., Mogel G.T., Phillips R., Robb R.A. and Vosburgh K.G. (Eds). Studies In Health Technology and Informatics Volume 132. Medicine Meets Virtual Reality 16 - parallel, combinatorial, convergent: NextMed by Design. pp. 446-50. 2008.

Moses GR and Park AE. "Ergonomic risk associated with assisting in minimally invasive surgery. " accepted in Westwood J.D., Haluck R.S., Hoffman H.M., Mogel G.T., Phillips R., Robb R.A. and Vosburgh K.G. (Eds). Studies In Health Technology and Informatics Volume 132. Medicine Meets Virtual Reality 16, 2008.

Dexter F, Xiao Y, Dow AJ, Strader MM, Ho D, Wachtel RE. "Coordination of Appointments for Anesthesia Care Outside of Operating Rooms Using an Enterprise Wide Scheduling System". *Anesthesia and Analgesia. 105:1701-1710.* 2007

Xiao Y, Schimpff S, Mackenzie CF, Merrell R, Entin E, Voigt R, Jarrell B. "Video Technology to Advance Safety in the Operating Room and Perioperative Environment". *Surgical Innovation. 14(1): 52-61. 2007*

Xiao Y, Dexter F, Hu FP, Dutton R. "Usage of Distributed Displays of Operating Room Video when Real-Time Occupancy Status was Available" . Anesthesia and Analgesia 2008; 106(2):554-560. 2008

Kim Y-J, Xiao Y, Hu P, Dutton RP. "Staff Acceptance of Video Monitoring for Coordination: A Video System to Support Perioperative Situation Awareness". *Journal of Clinical Nursing (accepted). 2007*

Dandekar O., K. Siddiqui, V. Walimbe, and R. Shekhar, "Image registration accuracy with low-dose CT: how low can we go?," in 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006, pp. 502-505.

Dandekar O., C. Castro-Pareja, and R. Shekhar, "FPGA-based real-time 3D image preprocessing for image-guided medical interventions," Journal of Real-Time Image Processing, vol. 1(4), pp. 285-301, 2007.

Shetye A.S. and R. Shekhar, "A statistical approach to high quality CT reconstruction at low radiation doses for real-time guidance and navigation," Proc. SPIE Med. Imaging, 2007.

Dandekar O., V. Walimbe, and R. Shekhar, "Hardware Implementation of Hierarchical Volume Subdivision-based Elastic Registration" in 28th Annual International Conference of the IEEE: Engineering in Medicine and Biology Society, 2006, pp. 1425-1428.

Dandekar O. and R. Shekhar, "FPGA-accelerated Deformable Registration for Improved Target-delineation During CT-guided Interventions," IEEE Transactions on Biomedical Circuits and Systems, vol. 1(2), pp. 116-127, 2007.

Dandekar O., W. Plishker, S. Bhattacharyya, and R. Shekahr, "Multiobjective Optimization of FPGA-Based Medical Image Registration" IEEE Symposium on Field-Programmable Custom Computing Machines, Under Review, 2008.

Shekhar R., O. Dandekar, S. Kavic, I. George, R. Mezrich, and A. Park, "Development of continuous CT-guided minimally invasive surgery," Multimedia Meets Virtual Reality (MMVR), 2007.

Shekhar R., O. Dandekar, S. Kavic, I. George, R. Mezrich, and A. Park, "Development of continuous CT-guided minimally invasive surgery," Proc SPIE, Medical Imaging 2007.

Lee G. and Adrian E. Park. "Development of a More Robust Tool for Postural Stability Analysis of Laparoscopic Surgeons". Surg Endosc (2008) 22:1087–1092

Lee G., T. Lee, D. Dexter, R. Klein, and A Park. "Methodological Infrastructure in Surgical Ergonomics: A Review of Tasks, Models, and Measurement Systems". Surgical Innovation, Volume 14 Number 3, September 2007 153-167

Lee G., S. M. Kavic, I. M. George and AE. Park. "Postural Instability Does Not Necessarily Correlate to Poor Performance: Case in Point". Received: 7 August 2006/Accepted: 22 September 2006/Online publication: 8 February 2007, Surg Endosc (2007) 21: 471–474

Lee G., T. Lee, D. Dexter, C. Godinez, N. Meenaghan, R.Catania and AE Park. "Ergonomic Risk Associated with Assisting in Minimally Invasive Surgery". Accepted August 2008, Surgical Endoscopy

APPENDICES

A. Slide Presentation: The Surgery Scheduling Problem, Block Release Policies, and Operations Research Applied to Health Care; William Herring, AMSC Ph.D. Preliminary Oral Exam University of Maryland, June 30, 2009, Dr. Jeffrey Herrmann, Chairman.

B. Ergonomic Safety of Surgical Techniques and Standing Positions Associated with Laparoscopic Cholecystectomy. Gyusung Lee, Yassar Youssef, Melody Carswell,
Cindy Hui-Lio, Ivan George and Adrian Park.

C. Laparoscopic Cholecystectomy Poses Physical Injury Risk to Surgeons: Analysis of Hand technique and Standing Position.   Yassar Youssef, Gyusung Lee, Carlos Godinez, Erica Sutton, Rosemary Klein, Ivan George, Jacob Seagull and Adrian Park.

D. A Ubiquitous Context-Aware Environment for Surgical Training.
P. Ordóñez, P. Kodeswaran, V. Korolev, W. Li, O. Walavalkar, B. Elgamil, A. Joshi, T. Finin, Y. Yesha.

E. Video Summarization of Laparoscopic Cholecystectomies. Michael A. Grasso, Timothy Finin, Anupam Joshi, and Yelena Yesha.

F. Cognitive Simulation in Virtual Patients. Sergei Nirenburg, Marjorie McShane, Stephen Beale, Bruce Jarrell and George Fantry.

G. Live Augmented Reality – A New Visualization Method for Paparoscopi SDurgery Using Continuous Volumetric CT.   Raj Shekhar, Venkatesh Bhat, Omkar Dandekar, Mathew Philip, Peng Lei, Carlos Godinez, Erica Sutton, Ivan George, Steven Kavic,
Reuben Mezrich, and Adrian Park.

H. An image registration–based approach for continuous volumetric CT   guided Interventions. Raj Shekhar and , Ananthranga Prithviraj.

I. Yang, R., Carswell, C.M., Wang, X., Zhang, Q., Han, Q., Lio, C., and Seales, B. Mapping the Way to a Dual Display Framework for Laparoscopic Surgery.

J. Lio, C.H., Carswell, C.M., Han, Q., Park, A., Strup, S., Selaes, W.B., Clarke, D., Lee, G., and Hoskins, J. (2009). Using Formal Qualitative Methods to Guide Early Development of an Augmented Reality Display System for Surgery. Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting. Santa Monica, CA: HFES.

K. Wang, X., Zhang, Q., Han, Q., Yang, R., Carswell, M., Seales, B., and Sutton, E. (2009) Endoscopic video texture mapping on pre-built 3D anatomical objects without camera tracking. IEEE Transactions on Medication Imaging, 7(7), 1-12.

L. Liao, M., Wang, L., Yang, R., and Gong, M. Real-time light fall-off stereo. (2008). International Conference on Image Processing (ICIP).

M. Han, Q. Strup, S., Carswell, C.M., Clarke, D., Seales, W.B. (2009). Model Completion via Deformation Cloning Based on an Explicit Global Deformation Model. 12th International Conference on Medical Image Computing & Computer Assisted Intervention (MICCAI)(1) 2009: 1067-1074.

N. A Research Portfolio for Innovation in the Surgical Environment. Gerald R. Moses and Adrian E. Park.

The Surgery Scheduling Problem, Block Release Policies, and Operations Research Applied to Health Care

Click to edit Master subtitle style

William Herring
AMSC Ph.D. Preliminary Oral Exam
University of Maryland
June 30, 2009

Dr. Jeffrey Herrmann, Chairman
Department of Mechanical Engineering
Institute for Systems Research

7/27/09

Overview

· Problem Statement & Literature Review

· Building Cyclic Master Surgery Schedules

(Beliën and Demuelemeester, 2007) [1]

· Operating Room Scheduling

(Guinet and Chaabane, 2003) [2]

· Releasing Allocated Block Time

(Dexter and Macario, 2004) [3]

· Conclusions

The Surgery Scheduling Problem

Underlying Problem

- The allocation of a fixed amount of resources under uncertain demand

- Resources: Operating rooms, Surgical equipment, Staff, Post-operative beds

- Uncertain Demand: Number of surgical cases, Duration of surgical cases, Length of post-operative stay for surgical cases

Motivation

- Operating room is most resource-intensive and profitable unit of a hospital

Three Scheduling Stages [1,4,10,11]

(1) Case Mix Planning
· Long-term
· Determining how much time to allocate to different surgical specialties

(2) Block Schedule
· Medium-term
· Determining which specialties get access to which operating rooms on which days

(3) Patient Scheduling
- Short-term
· Determining which individual patients to schedule on which days and in what order

Scheduling Objectives by Stage

(1)   Case Mix Planning

  - Meeting specialties' demand [11], Achieving throughput goals [10], Maximizing revenue [8]

(1)   Block Schedule

  - Leveling hospital bed occupancy [1], Minimizing overcapacity [12]

(1)   Patient Scheduling

  - Minimizing patient waiting times [2], Maximizing operating room efficiency [5]

Building a Block Schedule

"Building cyclic master surgery schedules with leveled resulting bed occupancy"

(Beliën and Demeulemeester, 2007) [1]

Objective: Construct a cyclic master surgery, or block, schedule that minimizes the total expected bed shortage (TEBS) over the length of the cycle

Schedule Requirements:
· Each surgeon (or specialty) requires a certain number of blocks per cycle
· There are a fixed number of OR blocks available on each day of the cycle

Assumptions:
· The number of patients operated on per block depends on the surgeon and is deterministic
· The length of stay for each patient depends on the surgeon and is stochastic following a multinomial distribution

Mathematical Formulation

$$\text{Min} \quad TEBS = \sum_{i \in A} EBS_i$$

$$\text{s.t.} \quad \sum_{i \in A} x_{is} = r_s \quad \forall \, s \in S$$

$$\sum_{s \in S} x_{is} \leq b_i \quad \forall \, i \in A$$

$$x_{is} \in \{0, 1, 2, \ldots, \min(r_s, b_i)\}$$

$$\forall \, s \in S \text{ and } \forall \, i \in A$$

Mathematical Formulation

Min $\quad TEBS = \sum_{i \in A} EBS_i$

s.t. $\quad \sum_{i \in A} x_{is} = r_s \quad \forall\, s \in S$

$\quad\quad \sum_{s \in S} x_{is} \leq b_i \quad \forall\, i \in A$

$\quad\quad x_{is} \in \{0,1,2,\ldots,\min(r_s, b_i)\}$

$\quad\quad\quad \forall\, s \in S \text{ and } \forall\, i \in A$

$U_{ijs} =$ beds occupied on day i resulting from

$\quad\quad\quad$ surgery on day j by specialty s

$Z_i =$ total number of beds occupied on day i

$\quad = \sum_{s \in S} \sum_{j \in A} U_{ijs} \quad$ (dependent on x_{is})

$EBS_i = E[\max(0, Z_i - c_i)] = \sum_{z_i = c_i + 1}^{\infty} (z_i - c_i) f_Z(z_i)$

$\quad$ where $c_i =$ bed capacity on day i

THE OBJECTIVE FUNCTION IS NONLINEAR IN THE DECISION VARIABLES!!

Solution Approaches

(1) Linearize the objective function and solve MIP

- Minimize the maximum expected bed capacity over the length of the cycle

⭐

- Minimize the maximum weighted combination of bed capacity mean and variance

- Minimize the maximum "mean + N*stdev" of bed capacity over the length of the cycle

(2) Heuristics with nonlinear objective functions

⭐

- Embed a MIP into an iterative heuristic that uses the Central Limit Theorem to approximate *TEBS*

⭐

- Define a quadratic MIP to level expected capacity peaks

- Use simulated annealing and approximate *TEBS* using CLT

Linearization

Claim: The mean and variance of Z_i are linear in x_{js}.

Mean:

$$\mu_i = E[Z_i] = \sum_{s \in S} \sum_{j \in A} E[U_{ijs}]$$

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d=\text{dist}(i,j)}^{m_s} E[D_{sd}] \lceil d/l \rceil \right) x_{js}$$

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d=\text{dist}(i,j)}^{m_s} p_{sd} n_s \lceil d/l \rceil \right) x_{js}$$

D_{sd} = number of patients in hospital d days after one block by surgeon s

$\sim binom(p_{sd}, n_s)$

Linearization

Claim: The mean and variance of Z_i are linear in x_{js}.

Mean: $\mu_i = E[Z_i] = \sum_{s \in S} \sum_{j \in A} E[U_{ijs}]$

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d=\text{dist}(i,j)}^{m_s} E[D_{sd}] \lceil d/l \rceil \right) x_{js}$$

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d=\text{dist}(i,j)}^{m_s} p_{sd} n_s \lceil d/l \rceil \right) x_{js}$$

D_{sd} = number of patients in hospital d days after one block by surgeon s

$\sim binom(p_{sd}, n_s)$

$\sigma_i^2 = Var[Z_i]$

Variance:

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d=\text{dist}(i,j)}^{m_s} p_{sd}(1-p_{sd})n_s \lceil d/l \rceil - \sum_{d_1=\text{dist}(i,j)}^{m_s} \sum_{d_2=\text{dist}(i,j)}^{d_1-1} 2 p_{sd_1} p_{sd_2} n_s \lceil d_2/l \rceil \right) x_{js}$$

Linearization

Claim: The mean and variance of Z_i are linear in x_{js}.

D_{sd} = number of patients in hospital d days after one block by surgeon s

$\sim binom(p_{sd}, n_s)$

Mean:

$$\mu_i = E[Z_i] = \sum_{s \in S} \sum_{j \in A} E[U_{ijs}]$$

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d = \text{dist}(i,j)}^{m_s} E[D_{sd}] \lceil d/l \rceil \right) x_{js}$$

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d = \text{dist}(i,j)}^{m_s} p_{sd} n_s \lceil d/l \rceil \right) x_{js}$$

$$\sigma_i^2 = Var[Z_i]$$

Variance of binomial r.v.

Covariance of 2 binomial r.v.'s

Variance:

$$= \sum_{s \in S} \sum_{j \in A} \left(\sum_{d = \text{dist}(i,j)}^{m_s} p_{sd}(1 - p_{sd}) n_s \lceil d/l \rceil - \sum_{d_1 = \text{dist}(i,j)}^{m_s} \sum_{d_2 = \text{dist}(i,j)}^{d_1 - 1} 2 p_{sd_1} p_{sd_2} n_s \lceil d_2/l \rceil \right) x_{js}$$

Mixed Integer Program Models

Min $\quad \gamma$

s.t. $\quad \displaystyle\sum_{i \in A} x_{is} = r_s \quad \forall\, s \in S$

$\displaystyle\sum_{s \in S} x_{is} \leq b_i \quad \forall\, i \in A$

$w_\mu \mu_i + w_{\sigma^2} \sigma_i^2 \leq \gamma \quad \forall\, i = 1,\ldots,l$

$x_{is} \in \{0,1,2,\ldots,\min(r_s,b_i)\}$

$\forall\, s \in S \text{ and } \forall\, i \in A$

Note: The weights for the mean and the variance can be modified to reflect the preferences of a decision-maker or to reflect the conditions in different hospital settings.

Direct Approximation of *TEBS*

By the Central Limit Theorem, Z_i approximately distributed as $N(\mu_i, \sigma_i^2)$

Therefore, $EBS_i \approx \displaystyle\int_{c_i + 0.5}^{\infty} (z_i - c_i) \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(z_i - \mu_i)^2}{2\sigma_i^2}} dz_i$

Resulting Solution Approaches:
· Use a heuristic approach to compute a feasible block schedule
· Evaluate the schedule based on this approximation of *TEBS*
· Heuristics: Build off earlier MIP's, Simulated Annealing

Iterative MIP Heuristic

1515

Iterative MIP Heuristic

Example of Iterative MIP Heuristic
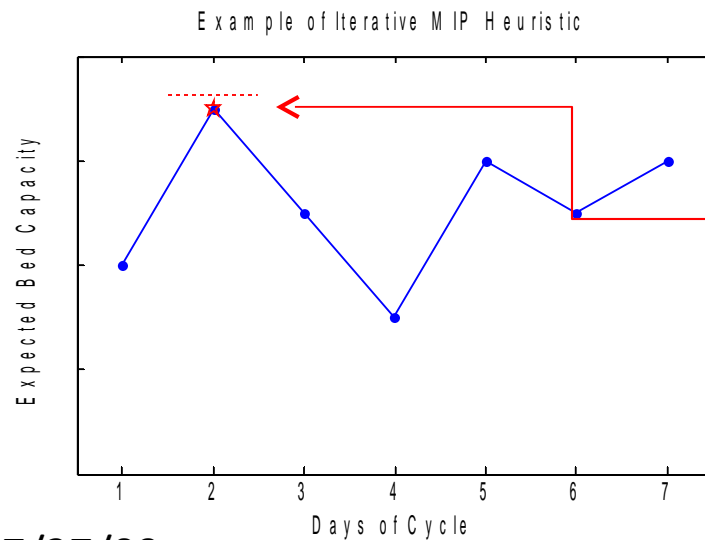
Day 2 has highest expected bed capacity

Place a hard constraint on μ_2

Minimize $\max\{\mu_1, \mu_3, \mu_4, \mu_5, \mu_6, \mu_7\}$

Simulated Annealing Heuristic

- Evaluate solutions by approximating *TEBS*

- Generating new solutions:

  - Choose a block at random

  - Create new schedule by swapping with another block

- Accepting new solutions:

  - Accept first swap that gives a smaller *TEBS*, $\exp(\Delta\,TEBS\,/\,T)$

  - If no swap yields improvement, choose swap with smallest increase and accept it with probability

Note: Each evaluation of a feasible swap requires full evaluation of *TEBS* via numerical integration, even if the swap ultimately isn't accepted

Testing and Results

Test Problems

· 384 test problems

· Vary a range of factors: Blocks per day, Number of specialties, Patients per specialty's block, Division of blocks, Range of LOS distribution, Probability of cancellation, and Available capacity

Results

· Simulated annealing found best solutions, but with poor computation time

· Pure MIP approaches found worst solutions, but did so very quickly

· Iterative MIP heuristics had the best combination of solution quality and computation time

· Most significant factors for solving difficulty: Blocks per day, Number of specialties, and to lesser degrees the Patients per block and LOS

Scheduling Individual Patients

"Operating theatre planning" (Guinet and Chaabane, 2003) [2]

Problem: Schedule a given number of patients into available operating room space over several days with constraints on case deadlines, resources, and surgeon availability.

Objective: Minimize the hospitalization costs of patients waiting for surgery plus the overtime costs from the operating rooms.
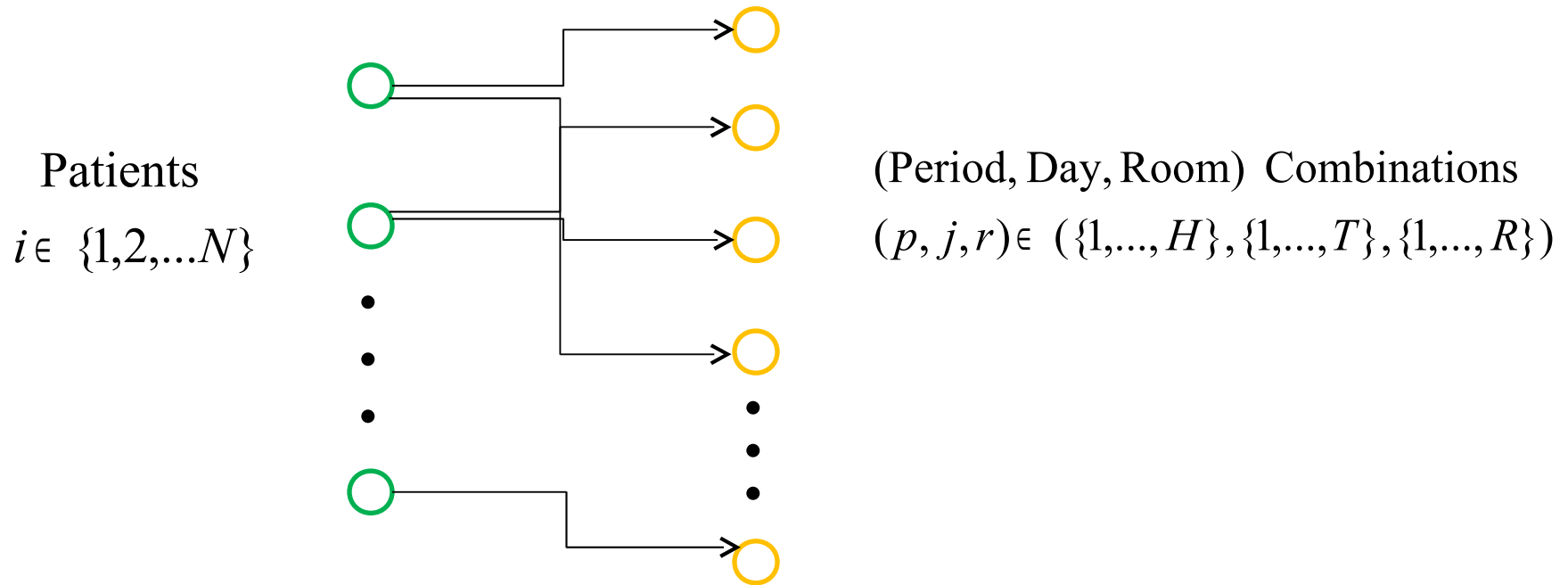
Schedule Requirements:
· Patients have a time window of a few days inside which to receive surgery
· Limitations on regular and overtime operating room time
· Limitations on surgeon availability (both days and hours per day)
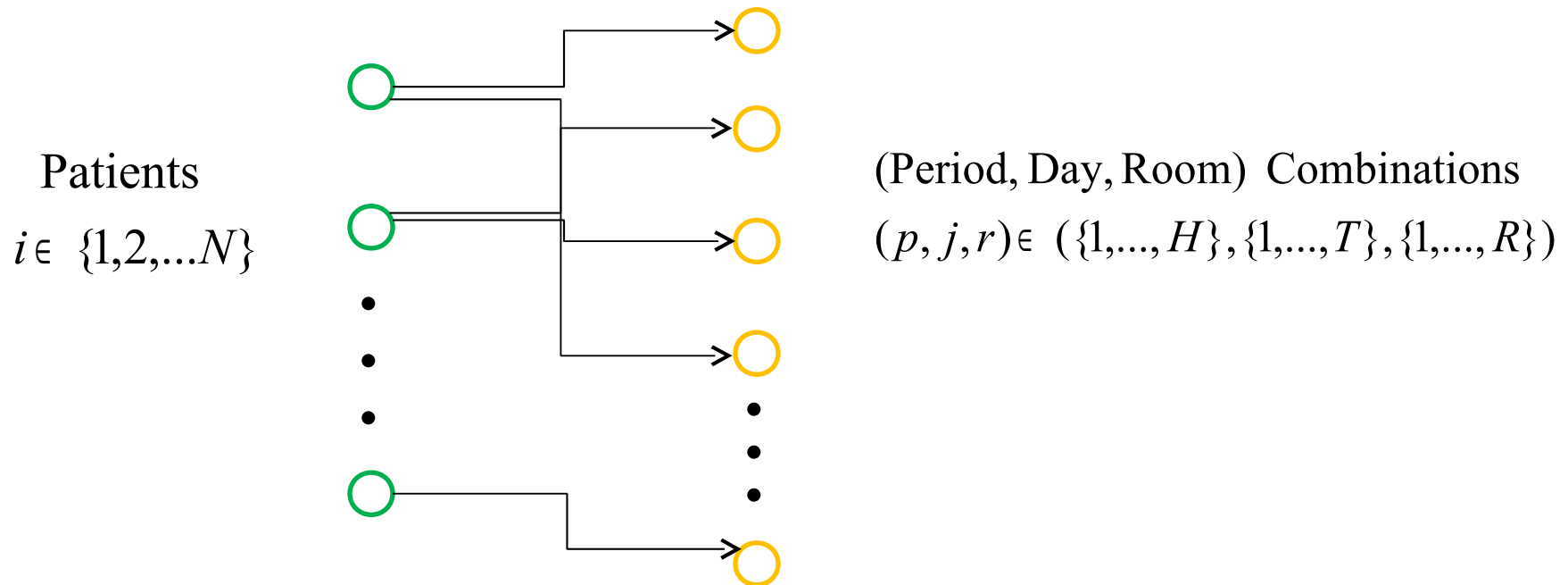· Limitations on equipment availability in certain rooms

Assumptions:
· Round predicted case durations to 1,2,3, or 4 hours
· Secondary resources such as recovery beds and carriers are sufficiently planned

Underlying Assignment Problem

Patients

$i \in \{1, 2, \ldots N\}$

(Period, Day, Room) Combinations

$(p, j, r) \in (\{1, \ldots, H\}, \{1, \ldots, T\}, \{1, \ldots, R\})$

Underlying Assignment Problem

Patients

$i \in \{1,2,...N\}$

(Period, Day, Room) Combinations
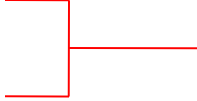
$(p, j, r) \in (\{1,...,H\}, \{1,...,T\}, \{1,...,R\})$

· Patient hospitalization date and deadline, as well as case duration, limit which arcs are feasible
· Arcs have a weight (case duration) as well as a cost
· Additional constraints use arc weights to enforce limits on OR and surgeon time availability
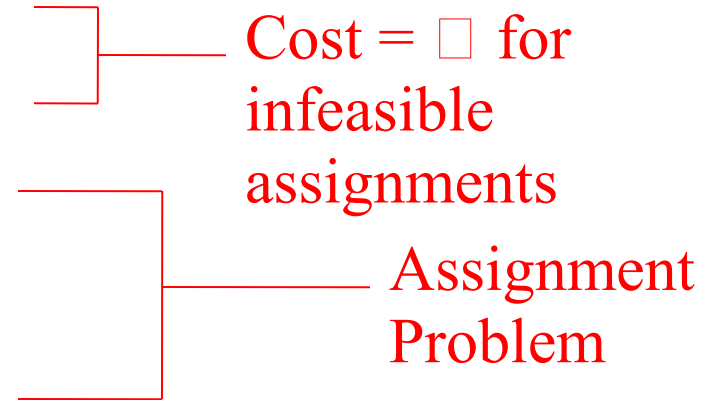· Goal is to find a complete assignment with minimal cost

Case Durations and Overtime

- The model does NOT attempt to sequence patients within rooms, rather the *periods* for each room in each day are used to handle different case durations and enforce overtime costs

- Consider a room with 8 regular and 4 overtime hours available, with the condition that cases over 2 hours are not allowed in overtime

- Possible case durations: 1, 2, 3, or 4

- Define 12 periods: Regular periods of length 1, 1, 1, 1, 2, 2, 4, 4

  Overtime periods of length 1, 1, 2, 2

- Any combination of cases filling the available hours can be fit into these periods, and overtime costs can directly be assigned to overtime periods
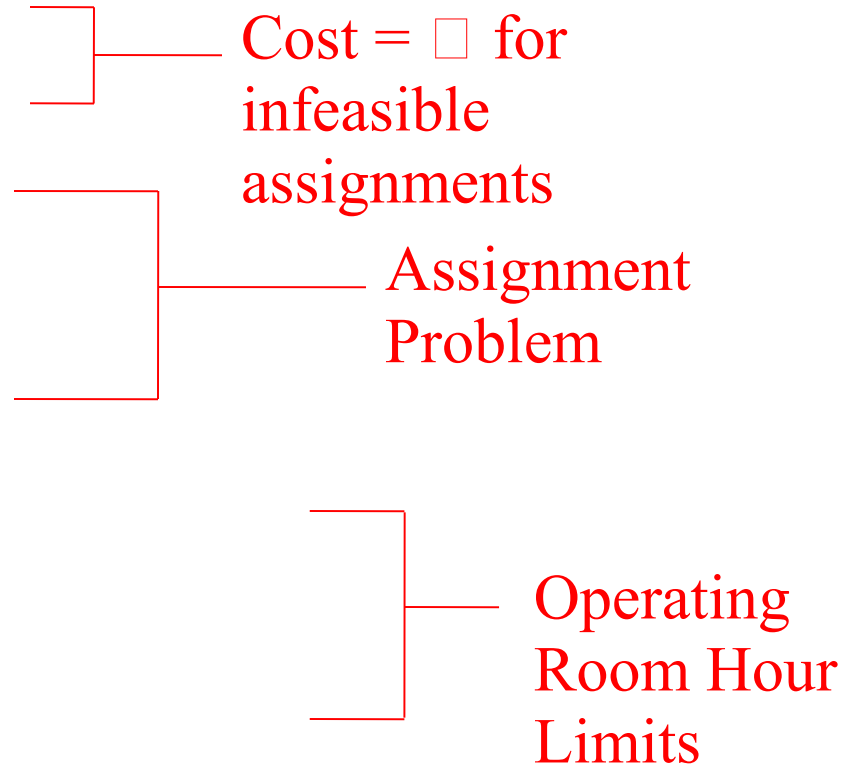
Formulation

Cost = □ for infeasible assignments

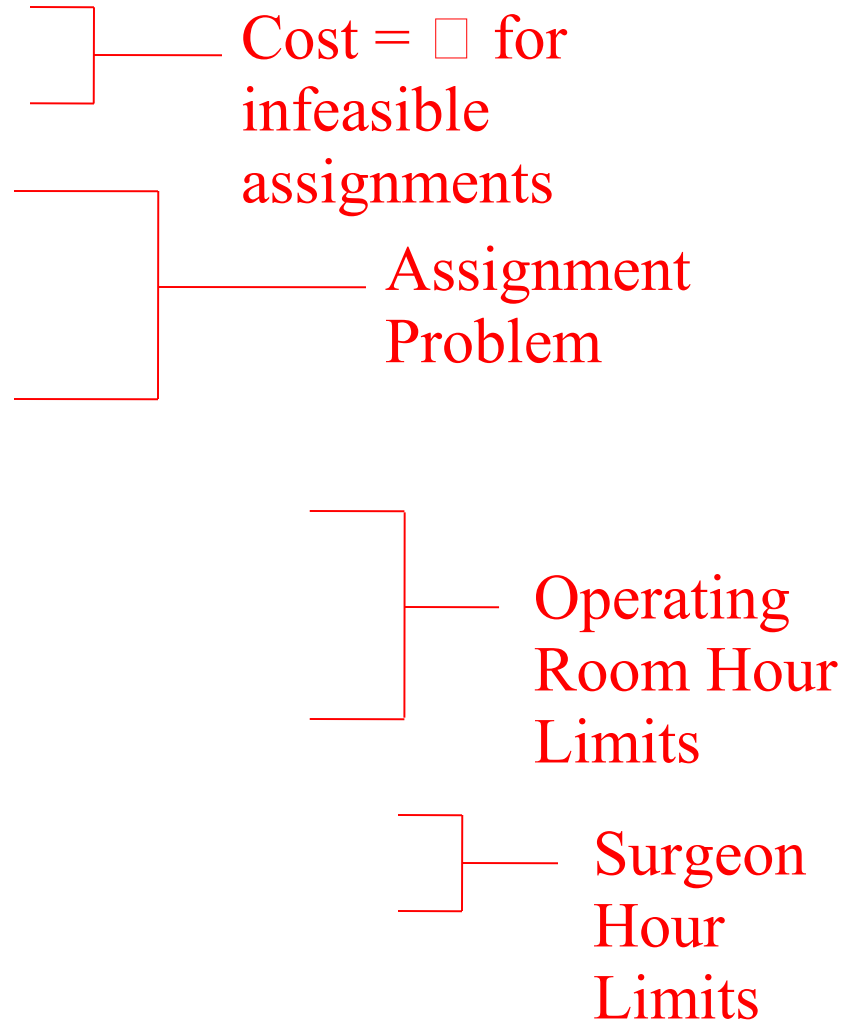# Formulation

Cost = □ for infeasible assignments

Assignment Problem

Formulation

Cost = □ for infeasible assignments

Assignment Problem

Operating Room Hour Limits

Formulation

Cost = □ for infeasible assignments

Assignment Problem

Operating Room Hour Limits

Surgeon Hour Limits

Primal-Dual Heuristic

Recall the Hungarian algorithm for the classical Assignment Problem [9]

· Use dual of LP relaxation to generate initial solution

· Use augmenting paths to increase size of matching in current solution

· If resulting matching is not complete, then find a new dual solution and iterate until a complete matching is found

Modifications for Surgery Scheduling

· Same dual structure is used on underlying assignment problem

· When finding augmenting paths, the feasibility (to *all* constraints) is checked before changing the solution

· Stops when a feasible solution that schedules all patients s found (no guarantee of optimality)

Data for Test Problems

- Number of days: $T=5$

- Number of cases: $N=10, 15, 20, \ldots, 85$

- Operating rooms: $M=1, 2, 3$

- Case durations have mean 2 and standard deviation 1 (hours)

- Hospitalization dates and surgery deadlines have means 2 and 4, respectively, and standard deviations of 1 (days of week)

- 8 regular hours and 4 overtime hours

- Patients can be scheduled with any surgeon (unrealistic in U.S.)

- Equipment requirements not considered

- 19 problem sizes, with 32 instances generated for each size

Test Results

Table 2
Ratio averages in percentage

No	N	M	Load	Planned	Mean $(((H-LB)/LB)100)$	Max $(((H-LB)/LB)100)$	Optimum
1	10	1	33.33	100.00	0.59	6.45	90.06
2	15	1	50.00	100.00	1.43	10.34	71.88
3	20	1	66.67	93.75	1.26	5.41	63.33
4	25	1	83.33	59.38	4.78	11.20	21.05
5	30	1	100.00	56.25	4.18	9.15	16.66
6	30	2	50.00	100.00	0.68	5.66	84.38
7	35	2	58.33	100.00	1.36	9.77	65.63
8	40	2	66.67	96.88	2.04	8.55	41.94
9	45	2	75.00	71.88	3.20	9.79	13.04
10	50	2	83.33	71.88	5.00	10.04	4.35
11	55	2	91.67	59.38	3.97	9.81	5.26
12	60	2	100.00	71.88	3.41	9.34	8.70
13	55	3	61.11	100.00	0.84	4.69	84.38
14	60	3	66.67	100.00	0.40	3.28	81.25
15	65	3	72.22	93.75	0.46	2.99	71.88
16	70	3	77.78	96.88	0.63	4.08	71.88
17	75	3	83.33	100.00	0.58	5.07	75.00
18	80	3	88.89	96.88	0.40	2.14	71.88
19	85	3	94.44	96.88	0.42	3.59	75.00
Averages			73.62	87.12	1.87	6.91	53.10

$$\text{Note}: Load = \frac{N \times 2}{M \times 60}$$

Test Results

Table 2
Ratio averages in percentage

No	N	M	Load	Planned	Mean $(((H-LB)/LB)100)$	Max $(((H-LB)/LB)100)$	Optimum
1	10	1	33.33	100.00	0.59	6.45	90.06
2	15	1	50.00	100.00	1.43	10.34	71.88
3	20	1	66.67	93.75	1.26	5.41	63.33
4	25	1	83.33	59.38	4.78	11.20	21.05
5	30	1	100.00	56.25	4.18	9.15	16.66
6	30	2	50.00	100.00	0.68	5.66	84.38
7	35	2	58.33	100.00	1.36	9.77	65.63
8	40	2	66.67	96.88	2.04	8.55	41.94
9	45	2	75.00	71.88	3.20	9.79	13.04
10	50	2	83.33	71.88	5.00	10.04	4.35
11	55	2	91.67	59.38	3.97	9.81	5.26
12	60	2	100.00	71.88	3.41	9.34	8.70
13	55	3	61.11	100.00	0.84	4.69	84.38
14	60	3	66.67	100.00	0.40	3.28	81.25
15	65	3	72.22	93.75	0.46	2.99	71.88
16	70	3	77.78	96.88	0.63	4.08	71.88
17	75	3	83.33	100.00	0.58	5.07	75.00
18	80	3	88.89	96.88	0.40	2.14	71.88
19	85	3	94.44	96.88	0.42	3.59	75.00
Averages			73.62	87.12	1.87	6.91	53.10

$$\text{Note}: Load = \frac{N \times 2}{M \times 60}$$

Test Results

Table 2
Ratio averages in percentage

No	N	M	Load	Planned	Mean $(((H-\text{LB})/\text{LB})100)$	Max $(((H-\text{LB})/\text{LB})100)$	Optimum
1	10	1	33.33	100.00	0.59	6.45	90.06
2	15	1	50.00	100.00	1.43	10.34	71.88
3	20	1	66.67	93.75	1.26	5.41	63.33
4	25	1	83.33	59.38	4.78	11.20	21.05
5	30	1	100.00	56.25	4.18	9.15	16.66
6	30	2	50.00	100.00	0.68	5.66	84.38
7	35	2	58.33	100.00	1.36	9.77	65.63
8	40	2	66.67	96.88	2.04	8.55	41.94
9	45	2	75.00	71.88	3.20	9.79	13.04
10	50	2	83.33	71.88	5.00	10.04	4.35
11	55	2	91.67	59.38	3.97	9.81	5.26
12	60	2	100.00	71.88	3.41	9.34	8.70
13	55	3	61.11	100.00	0.84	4.69	84.38
14	60	3	66.67	100.00	0.40	3.28	81.25
15	65	3	72.22	93.75	0.46	2.99	71.88
16	70	3	77.78	96.88	0.63	4.08	71.88
17	75	3	83.33	100.00	0.58	5.07	75.00
18	80	3	88.89	96.88	0.40	2.14	71.88
19	85	3	94.44	96.88	0.42	3.59	75.00
Averages			73.62	87.12	1.87	6.91	53.10

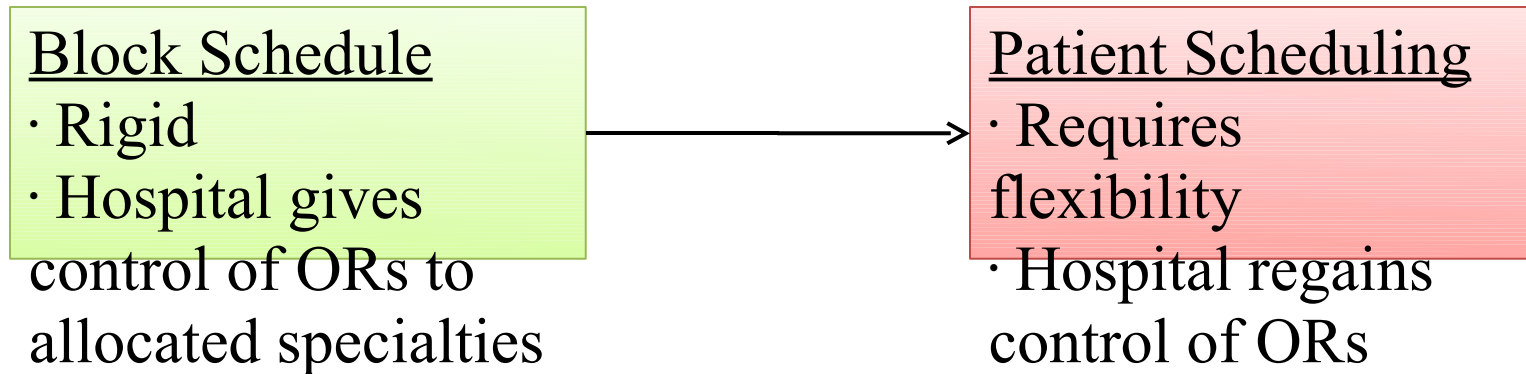Note : $Load = \dfrac{N \times 2}{M \times 60}$

Comments

Solving Difficulties
· Problem instances with Loads over 80% became much more difficult to schedule completely
· In each case, these difficulties were alleviated by adding an additional operating room
· However, these difficulties lessened with the larger problems, reflecting the larger number of available periods

Quality of Found Solutions
· The average distance from the lower bound (solution to just the underlying AP) is quite small (<2%)
· When a feasible solution is found, the solution is optimal (by comparison to underlying AP) about half the time
· The authors don't report on the size of the infeasibility for the unsolved problems (i.e. how many patients are left unscheduled)

An Understudied Interaction

How does the block schedule interact with patient scheduling?

Block Schedule
· Rigid
· Hospital gives control of ORs to allocated specialties

Patient Scheduling
· Requires flexibility
· Hospital regains control of ORs

An Understudied Interaction

How does the block schedule interact with patient scheduling?

Block Schedule
· Rigid
· Hospital gives control of ORs to allocated specialties
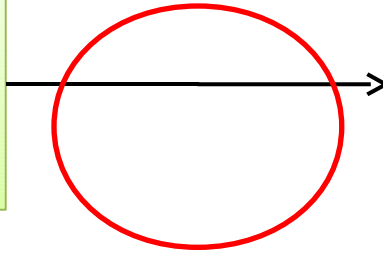
Patient Scheduling
· Requires flexibility
· Hospital regains control of ORs

Key Questions:
· What has to happen here for this to work?
· In what scenarios is this important?

Comments:
· When demand for a specialty doesn't always fill up its allocated time (common in U.S.), the unused time could be better used elsewhere.

3434

Existing Work on Block Release Policies

"When to release allocated operating room time to increase operating room efficiency"

(Dexter and Macario, 2004) [3]

<u>Conclusions from Prior Work:</u>
· Allocated time should only be released if there is a case waiting for it [6]
· The choice of which block to release should be made based on which room is *expected* to have the most unused time on the day of surgery [7]
· However, there is little gain in efficiency by releasing the room that has the most unused time *at the time of the release* [7]

<u>Question for Current Work:</u>
· Assuming that an add-on case is waiting for operating room time, how far ahead of the day of surgery should the block release take place?

Experiment Design

Use of Real Hospital Data

- 3 years of regularly scheduled OR days (including *when* cases were added to schedule), amounting to 754 test schedules

- Cases were fit into a hypothetical block schedule that was selected to maximize OR efficiency

Simulating Add-On Cases

- One hypothetical case was generated for each OR day, varying in length from 1 to 3 hours

- Block releases were considered on 1, 3, and 5 days before the day of surgery, and new cases were placed according to earlier findings

- Remaining cases (arriving after block release) were left in their originally scheduled rooms

- Operating room efficiency computed based on final day's schedule for

Findings and Shortcomings

Results

· Findings indicate that the timing of the block release has little impact on resulting overtime costs (<15 minutes/day).

· As a result, authors suggest that hospitals choose their block release policies based on their own staff's preferences.

Areas for Exploration

- The authors claim that multiple add-on cases can be handled one-by-one in the order of their arrival. Is this really optimal?
- Also, in a realistic setting, the placement of add-on cases into released block time can substantially change the scheduling decisions made after the block release. The authors don't consider this.

Questions?

References

Primary Source Material

[1]   Beliën, J., Demeulemeester E. 2007. Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research*, 176:1185-1204.

[2]   Guinet A., Chaabane, S. 2003. Operating theatre planning. *International Journal of Production Economics*, 85:69-81.

[3]   Dexter, F., Macario, A. 2004. When to release allocated operating room time to increase operating room efficiency. *Anesthesia & Analgesia*, 98:758-762.

Secondary Source Material

[4]   Blake, J.T., Donald, J. 2002. Mount Sinai Hospital uses integer programming to allocate operating room time. *Interfaces*, 32(2):63-73.

[5]   Denton, B., Viapiano, J., Vogl, A. 2007. Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science*, 10:13-24.

[6]   Dexter, F., Traub, R.D. 2002. How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. *Anesthesia & Analgesia*, 94: 933-942.

[7]   Dexter, F., Traub, R.D., Macario, A. 2003. How to release allocated operating room time to increase efficiency: predicting which surgical service will have the most under-utilized operating room time. *Anesthesia & Analgesia*, 96:507-512.

Overview of Research Project

Goal:

Develop a patient-flow model for the surgery scheduling process for a single day in the OR in order to explore block release and request queue policies in a more general setting.
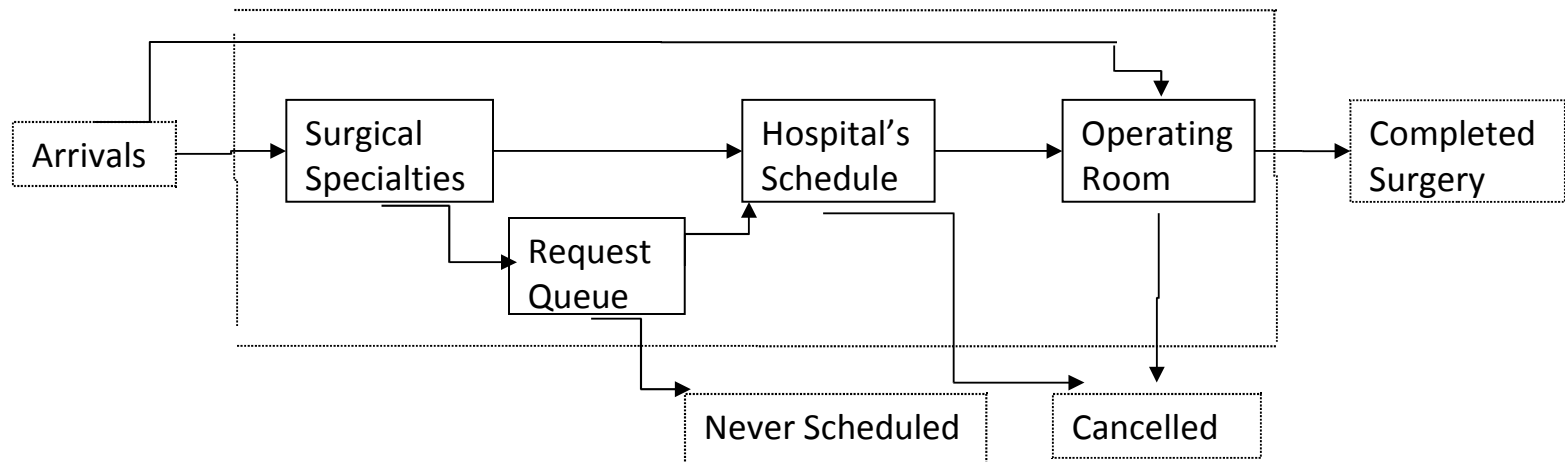
Ergonomic Safety of Surgical Techniques and Standing Positions Associated with Laparoscopic Cholecystectomy

Gyusung Lee[1], PhD, Yassar Youssef[2], MD, Melody Carswell[3], PhD
Cindy Hui-Lio[3], PhD, Ivan George[1], Adrian Park[1], MD

1. Department of Surgery, University of Maryland, Baltimore, MD, USA
2. Department of Surgery, Sinai Hospital of Baltimore, MD, USA
3. Department of Psychology, Center for Visualization and Virtual Environments, University of Kentucky, Lexington, KY, USA

Laparoscopic cholecystectomy (LC), a procedure in which, using either a one-handed or two-handed technique, a surgeon removes a symptomatic gallbladder in a minimally invasive manner, is commonly—due to its relatively high safety level—the initial procedure that a resident will perform. Investigation of the ergonomics associated with LC one-handed and two-handed techniques is one goal of this study. Identification of which of two standing positions (between legs or at side) used during LC is the more ergonomically favorable is the other. Knowledge gained from our research in these issues is intended to be applicable both to surgical training and the operating room environment. Eight right-handed laparoscopic surgeons with varying levels of surgical skills were recruited for this study. Each performed LC a total of four times on a virtual reality (VR) simulator with each performance incorporating one of the following conditions: either the one-handed or two-handed surgical technique or the position of standing between the patient's legs or at the patient's side. Each trial was also divided into two phases: 1) dissection and clipping and 2) gall bladder removal. During the performance of LC, physical ergonomic data were collected though surface electrode electromyography (EMG) and two force plates. Additionally NASA-Task Load Index (TLX) and secondary time estimation were used for cognitive ergonomic assessment. Standing at the side produced a significantly higher weight-loading ratio (WLR) than standing between the legs. Comparison between techniques indicated that the two-handed technique caused higher WLR. Significant phase effect equated increased WLR with phase 2 gall bladder removal. No statistical interactions among technique, standing position, and phase were significant. Analysis of NASA-TLX showed that global workload, influenced mainly by significant physical workload and effort scales, was higher with the side-standing position and the two-handed technique. The results from time estimation analysis, although statistically marginal, demonstrated that the one-handed technique is more mentally demanding. Our study demonstrated that due to lower physical as well as mental workload, the two-handed technique performed with the surgeon positioned between the patient's legs is the most ergonomically favorable combination. Additionally, it was demonstrated that the pedal for cautery operation requires ergonomic improvement. These specific findings encourage us to continue research into what proof ergonomics can provide regarding what constitutes the most efficacious approaches to surgical procedures and to optimizing patient safety and the surgical environment.

INTRODUCTION

Laparoscopic cholecystectomy (LC), currently one of the most performed minimally invasive surgical procedures in general surgery, has become the gold standard for the treatment of symptomatic cholelithiasis (gallstones) (Mosimann et al., 2006, Sain et al., 1996). LC is considered very safe with a morbidity of less than 5% and a mortality of less than 0.1% (Mosimann et al., 2006, Sain et al., 1996). Both the high incidence of LC being performed and its relative safety compared to other laparoscopic procedures make it ideal as the initial procedure to have residents perform in a real operating room following laboratory training.

Laparoscopic surgery is considered a two-handed technique, one in which the surgeon performing the procedure uses both hands in a kinetic manner. Thus, it is assumed that surgeon performing LC will use the two-handed technique, allowing the surgeon to manipulate instruments with both hands for retraction, dissection, clipping, transection, and gallbladder retrieval. Specifically, the surgeon will continuously retract the gallbladder neck with the left hand while dissecting the cystic duct and artery with the right one. However, the assumption that the bimanual technique will be used is inaccurate. At many institutions the surgeon dissect the gallbladder with one hand and hold the camera with the other while the assistant holds the gallbladder cephalad with the left hand and retracts the gallbladder with his/her right hand so that the area of dissection is exposed for the main surgeon.

Patterns and dynamics associated with ergonomics and performance may be discerned through analyzing data acquired in regard to two significant factors: technique and standing position. In addition to such investigation, the present study serves as a continuation of our previous work into the ergonomics governing surgical postural stability. That work

included a series of studies investigating postural control strategies used by surgeons of differing experience levels as they performed laparoscopic tasks. In two studies analyses of sway amplitude and sway area demonstrated that more experienced surgeons used unique postural controls during each task and their strategies differed from those used by less experienced surgeons (Lee et al., 2006, 2008). A third study demonstrated that postural instability does not necessarily correlate to poor performance as in the case of a surgeon using particular compensatory and/or strategic upper body movements to achieve optimal performance (Lee et al., 2007). Recent research investigated the ergonomics associated with the MIS surgical assistant (Lee et al., 2008). That study showed that during a simulated Nissen fundoplication, surgical assistants performing camera pointing and tissue retraction tasks bore 70% to 80% of their whole body weight on their supporting leg, a high-risk ergonomics situation attributable to their leaning posture and extended arm. The current study continues our previous research regarding physical workload and represents our initial use of electromyography (EMG) assessment. Additionally, cognitive workload assessment—as evaluated through NASA-TLX and time estimation—is brought to bear on our present-day ergonomic study of LC techniques and standing position.

Through ergonomic assessment we aim to delineate the pros and cons of the two-handed vs. one-handed technique for performing laparoscopic tasks as well as to compare the effects of surgeon position (on the left side of the patient vs. in between the patient's legs) on performance. Analyses of posture, physical workload, and cognitive workload provide unique knowledge, particularly in terms of LC, permitting us to determine ergonomically favorable conditions both for training and real-time performance.

METHODS

This IRB-approved study was conducted at the Surgical Ergonomics Laboratory at the Maryland Advanced Simulation, Training, Research, and Innovation (MASTRI) Center, University of Maryland, School of Medicine (UMSOM). Eight, right-handed subjects possessing different levels of MIS experience were recruited for this study from the Departments of Surgery at UMSOM and at Sinai Hospital of Baltimore. Each subject performed LC on a LapVR[TM] virtual reality (VR) surgical simulator (Immersion Medical, Gaithersburg, MD). Prior to beginning surgical tasks, each subject had 12 surface electrodes (Delsys[TM], Boston, MA) attached to different muscle groups, including biceps, triceps, deltoid, trapezius, wrist flexor, and extensor at both upper extremities so that EMG signals could be recorded. As a reference for normalization, which permits the comparison of activation levels between different muscle groups, maximum voluntary contraction (MVC) levels of each muscle group were recorded for several seconds. All EMG data was collected at 1000Hz. These data were full-wave rectified and then filtered using a second-order Butterworth low-pass filter with cutoff frequency of 10Hz. The EMG data collected during surgical tasks were further processed; dividing them by MVC levels collected prior to each task allowed the data to be

shown as %MVC. After this normalization process, the time-integral of data over performance time was taken to calculate what we termed "relative muscular workload" (RMW) over the period of performance time. RMW gets higher with either a high level of muscle contraction during a short activation duration or a long activation duration even with a relatively low contraction level.

While performing LC, each subject stood on two AMTI[TM] force plates (Advanced Mechanical Technology Inc., Watertown, MA) with a leg on each plate. These plates measured at 200 Hz the amplitudes of the vertical ground reaction forces (VGRF) exerted by each leg onto each force plate. To quantify the balancing taking place between the two legs, we derived a weight-loading ratio (WLR) by dividing the left force plate VGRF by the total VGRF of both plates.

$$WLR = \frac{\text{Left VGRF}}{\text{Left VGRF} + \text{Right VGRF}} \times 100$$

Assessment of mental workload was achieved through two methods: the NASA-Task Load Index (NASA-TLX) and time-estimation. The NASA-TLX required participants rate their experienced level of workload along six scales (mental demand, physical demand, temporal demand, effort, performance, and frustration) with 0 for low and 100 for high. Time-estimation, as a secondary task, required that throughout task performance each participant say "Time" whenever he/she thought that 21 seconds had elapsed. The mean interval lengths per trial as well as the standard deviation of the intervals were used as workload metrics with the assumption being that when workload increases (and spare attention capacity decreases), intervals will become longer and more variable. Both data analysis and interpretation of NASA-TLX and time-estimation were performed at the University of Kentucky.

Each subject was required to perform the LC procedure four times. Each performance was governed by one of the following specific conditions (Figure 1): one-handed or two-handed surgical technique (technique effect) or standing between the patient's legs or standing at the patient's side (standing effect). The order of these conditions was randomized for each subject. During the procedure performance, a camcorder was directed so as to record what appeared on the virtual simulator screen while another camera recorded an external view of the participant's upper body movements. These video images permitted identification of what surgical tasks were performed during a specific time window. Data gathered in this manner was categorized into two phases: 1) dissection and clipping and 2) gall bladder removal, the latter achieved by use of an L-shaped hook cautery that was operated by a pedal at the surgeon's right foot.

When the two-handed technique was used, each subject manipulated one instrument with the right hand (performing dissection, clipping, transection, and gall bladder removal) and another instrument with the left hand (performing tissue retraction) as an assistant manipulated a camera at the subject's request. With the one-handed technique, the participating subject performed the same tasks with the right-

hand as with the two-handed technique and with the left hand performed camera navigation while the assistant performed tissue retraction according to the subject's instructions.

Statistical Analysis

An overall $2 \times 2 \times 2$ (technique × standing × phase) analysis of variance (ANOVA) with repeated measures was applied to all data to investigate the physical and cognitive ergonomics associated with different surgical techniques, standing positions, and procedural phases. Then the main effects of these three factors and their interactions were analyzed. The significance level was set at a p value of 0.05.

RESULTS

Physical Ergonomics

Weight-loading ratio (WLR). WLRs as evidenced among the two surgical techniques and two standing positions are shown in Figure 2. Standing positions produced the most noticeable WLR results. WLR was significantly higher when standing at the side than when standing between legs (M_{side}=82.05; $M_{between}$=52.10; $F(1,6)$=150.71; p<0.05; $\eta^2_{partial}$=.96). The leaning posture that subjects exhibited during the task while standing at the side caused their left leg to bear more than 80% of their whole body weight. In terms of the two different surgical techniques, WLR proved higher with the two-handed than with the one-handed ($M_{one-handed}$=63.95; $M_{two-handed}$=70.20; $F(1,6)$=46.43; p<0.05; $\eta^2_{partial}$=.87). Interestingly, significant main effect of phase showed WLR during phase 2 as being significantly higher than WLR during phase 1 ($F(1,6)$=31.90; p<0.05; $\eta^2_{partial}$=.84). This indicates that operating the pedal for cautery caused an unbalanced posture that was not ergonomically favorable regardless of techniques or standing positions. Among technique, standing position, and phase, there were no statistically significant interactions, which meant that the effect size associated with each did not significantly depend with the others.

Relative muscular workload (RMW). Means of RMW at different techniques and standing positions are summarized in Table 1. Higher RMWs were observed from the biceps, deltoid, and trapezius at the left arm when LC was performed by surgeons using the two-handed technique (p<.05). When the surgeon stood at the patient's side, significantly higher RMW was shown in the left deltoid muscle (p<.05). This was also the case with both the left and right forearm flexors though the effect was marginal (p<.08).

Performance time. The means and standard deviations of LC performance time as evidenced with each of the four varying conditions are summarized in Table 2. Times associated with the two-handed technique and with standing between legs were shorter, though not statistically significant (p>0.05).

Cognitive Ergonomics

Subjective workload. The overall level of workload experienced by the participants was calculated from the raw (unweighted) means of the six subscale scores of the NASA-TLX. A main effect in regard to standing position was that side standing was determined to be more demanding (M_{side}=55.57; $M_{between}$=42.08; $F(1,6)$=15.46; p <.05; $\eta^2_{partial}$ = .72). Data also disclosed a reliable interaction of position and hand use ($F(1,6)$=8.72; p <.05; $\eta^2_{partial}$ = .59); thus, the size of the position effect depended on whether the surgeon manipulated both instruments or only one. As Figure 3 and follow-up comparisons revealed, the standing position effect on workload was exacerbated when the surgeons performed LC using the two-handed technique.

Further analysis of each of the six NASA-TLX subscales revealed that the global effects described above were mainly due to perceived differences in physical load (M_{side}=60.71; $M_{between}$=35.0; $F(1,6)$=22.97; p<.005; $\eta^2_{partial}$=.79) and effort (M_{side}=60.71; $M_{between}$=33.91; $F(1,6)$=44.06; p<.005; $\eta^2_{partial}$=.88). The physical load and effort subscales also revealed the same interaction pattern illustrated by the overall NASA-TLX score between the surgeon's body position and use of one-handed or two-handed technique ($F_{physical}(1,6)$=6.57; p<.043; $\eta^2_{partial}$=.52; $F_{effort}(1,6)$=6.17; p<.048; $\eta^2_{partial}$=.51). There was also a main effect for technique for the physical demand subscale, with use of the two-handed technique associated with greater perceived workload ($M_{one-handed}$=41.07; $M_{two-handed}$=54.64; $F(1,6)$=5.74; p=.05; $\eta^2_{partial}$=.49).

Secondary task performance. Indication of 21-second intervals was used as a secondary task assessment of workload throughout the duration of each trial. Analysis derived from the standard deviation of interval lengths revealed a marginal effect of technique ($M_{one-handed}$=25.97; $M_{two-handed}$=13.81; $F(1,6)$=3.99; p=.069; $\eta^2_{partial}$=.40) that demonstrated the one-handed technique compared to the two-handed technique required more mental capacity for the task performance.

DISCUSSION

A review of the literature did not reveal studies comparing the ergonomics associated with the two-handed versus the one-handed technique in laparoscopy despite that many experts believe and echo that laparoscopy is a two-handed art. The reason why some surgeons use the one-handed technique in performing LC is not clear. But we surmise several reasons: surgeon preference (Hamilton et al., 2002) or expertise, ergonomic position more favorable for the surgeon (not having arm extended around the patient in order to access a trocar), or better environment in which to mentor the resident during surgery performance.

Using both physical and cognitive workload assessments, we were able to identify a number of ergonomic issues in relation to surgical techniques and standing positions used in LC performance. The physical workload analyses demonstrated a smaller WLR, representing a less unbalanced standing posture, with both the one-handed technique and the position of standing between the patient's legs, thus proving each to be more ergonomically favorable than its respective alternative. This finding was also supported by EMG analysis.

As traditional EMG analysis uses variables only capable of explaining muscle activation over a short period of time, we designed a new variable—RMW—to quantify muscular workload over a longer time period. During the one-handed technique, for instance, lower RMW indicated that, not surprisingly, the surgeon's left arm was less elevated. That lower elevation also proved to be the case when the surgeon stood between the patient's legs.

Proof of significant phase effect highlighted another ergonomic issue – the leaning posture exhibited by surgeons to operate the cautery pedal. This awkward standing posture may be improved by adding an on/off switch to the cautery instrument or by modifying the design of the pedal. The one-handed technique and the position of standing between the patient's legs also were proven through NASA-TLX cognitive workload analysis to necessitate lower levels of physical demand and effort.

Our results showed that the most ergonomically favorable combination is offered by using the one-handed technique and standing between the patient's legs. How realistic would it be, however, to perform LC using the one-handed technique while standing between a patient's legs? Most LC procedures in the US are performed with the patient in reverse Trendelenburg position (the body is laid flat on the back with the head higher than the feet) with the right side slightly up. The surgeon stands on the left side of the patient with the assistant on the right side. LC procedures may also be performed with the patient in a lithotomy position (laid on the back with knees that are bent, positioned above the hips, and spread apart by stirrups) and in a reverse Trendelenburg with the surgeon standing between the patient's legs while the assistant is on the right side. When the surgeon is standing on the patient's left side, either the two-handed or one-handed technique may be used. When the surgeon is positioned between the patient's legs, the LC performed is usually two-handed because of the ease and proximity of the trocars. Given such operative realities, it seems that the most useful ergonomic recommendation to be made about performing LC would be that the procedure be executed using the two-handed technique while standing between the patient's legs.

Regarding cognitive ergonomics, time-estimation analysis showed by marginal effect that the one-handed technique required more mental capacity. Of note is that this result is inconsistent with that of the NASA-TLX which showed higher physical demand and effort exerted with the two-handed technique. However, the time-estimation analysis result is consistent with the trends found in terms of both interval durations (longer for the one-handed condition) and perceived mental demand and frustration (slightly higher for the one-handed condition). It must be said though that given the relatively large effect size and the low number of participants, time-estimation analysis warrants further investigation. It may be the case, as has been demonstrated in other situations (Lio et al., 2007), that interval production is sensitive to increased cognitive demands related to prediction. We surmise that increased cognitive demands associated with the one-handed technique occur because the surgeon is providing instructions to the assistant performing tissue retraction while also accommodating the assistant's performance so as to complete LC procedure.

Surgical ergonomics, still considered a new field of ergonomic studies, has been expanding its research in laparoscopic surgery to include newly developed surgical techniques and technologies such as robotic surgery and Natural Orifice Translumenal Endoscopic Surgery (NOTES). Such expansion, while both notable and necessary, should not obscure the fact that there remain unexplored or less explored territories in laparoscopy. The ergonomics associated with the surgical training and performance of LC constitutes just such an area. Decisions regarding the technique and standing position used by surgeons during procedures are often formed based on experience and preference. Surgical outcomes also influence such decisions. That ergonomic factors, too often overlooked, are vital elements that should influence the choice of technique and standing position is clearly demonstrated by our research. Knowing that ergonomically assessing physical and mental workloads, as highlighted by our study, can undoubtedly result in improved training conditions as well as patient and surgeon safety in the operating room, we plan to use both subjective and objective ergonomic assessment tools to continue to investigate work efforts associated with the techniques, positions, and cognition commonly used in laparoscopic procedures.

ACKNOWLEDGMENTS

REFERENCES

Hamilton EC, Scott DJ, Fleming JB, Rege RV, Laycock PC, Bergen PC, et al. (2002) Comparison of video trainer and virtual reality training systems on acquisition of laparoscopic skills. *Surgical Endoscopy*, 16(3), 406-411.

Lee, G., Kavic, S. M., George, I. M., & Park, A. E. (2007). Postural instability does not necessarily correlate to poor performance: Case in point. *Surgical Endoscopy*, 21(3), 471-474.

Lee, G., Lee, T. H., Dexter, D. J., Godinez, C., Meenaghan, N., Catania, R., et al. (2009). Ergonomic risk of assisting in minimally invasive surgery. *Surgical Endoscopy*, 23(1), 182-188.

Lee, G., & Park, A. E. (2008). Development of a more robust tool for postural stability analysis of laparoscopic surgeons. *Surgical Endoscopy*. 22(4), 1087-1092.

Lee, G., Weiner, M., Kavic, S. M., George, I. M., & Park, A. E. (2006). Pilot study—Correlation between postural stability and performance time during fundamentals of laparoscopic surgery (FLS) tasks. *British Journal of Surgery*, 93(Suppl 1), 206.

Lio, C. H., Carswell, C. M., Seales, W. B., Clarke, D., Kurs, Y., & DeCuir, J. (2007). Using a global implicit measure of situation awareness during performance of laparascopic training tasks. 51st Annual Meeting of the Human Factors and Ergonomics Society (2007, Santa Monica, CA). Proceedings of Human Factors and Ergonomics Society, 1280-1282.

Mosimann, F. (2006). Laparoscopic cholecystectomy has become the new gold standard for the management of symptomatic gallbladder stones. Hepatogastroenterology, 2006;53(69), 1.

Sain, A. H. (1996). Laparoscopic cholecystectomy is the current ''gold standard'' for the treatment of gallstone disease. *Annals of Surgery*, 224(5), 689–690.
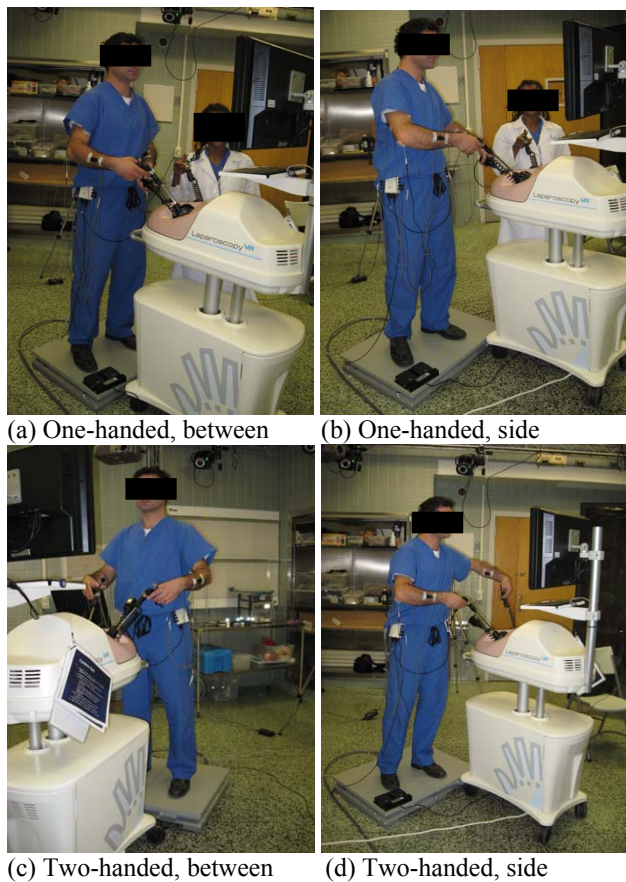
(a) One-handed, between

(b) One-handed, side

(c) Two-handed, between

(d) Two-handed, side

Figure 1. Experimental set-up for each condition.
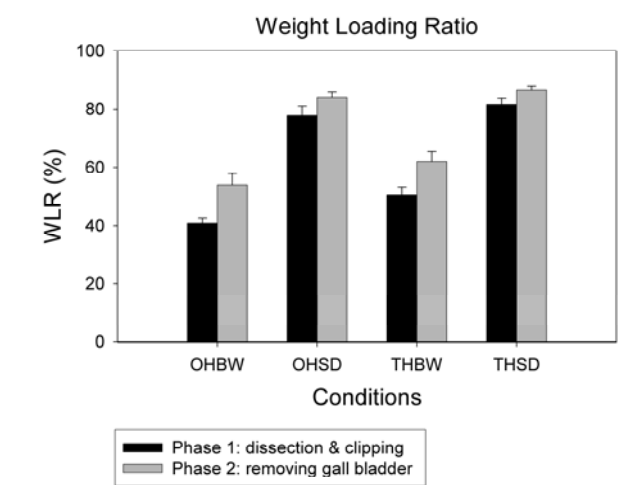
Figure 2. The weight-loading ratio (WLR) with four different conditions (OHBW: one-handed with between leg standing,

OHSD: one-handed with patient's side standing, THBW: two-handed with between leg standing, THSD: two-handed with patient's side standing)

Table 1. Summary of the relative muscular workload (RMW). Separate statistical comparison was applied to techniques and to standing positions with the pair of bolded numbers representing statistical difference (significant main effect, $p < .05$).

Muscles	Side	Techniques		Standing positions	
		One-handed	Two-handed	Between legs	Patient Side
Biceps	Right	1718.1	1639.8	1592.7	1765.2
	Left	**1443.9**	**2328.9**	1618.0	2154.8
Triceps	Right	1617.5	1574.8	1407.4	1784.9
	Left	3135.9	3424.9	2793.5	3767.4
Deltoid	Right	1298.9	1042.8	1114.2	1227.5
	Left	**1044.6**	**2481.8**	**959.6**	**2566.8**
Trapezius	Right	4692.8	3981.5	4417.9	4256.4
	Left	**5742.8**	**7732.0**	4062.5	9412.3
Forearm Flexor	Right	2533.3	2357.7	2188.3	2702.7
	Left	1525.2	1601.3	1298.5	1828.5
Forearm Extensor	Right	10857.5	9344.7	8526.9	11675.3
	Left	7223.7	6127.6	5815.6	7535.4

Table 2. Summary of performance time

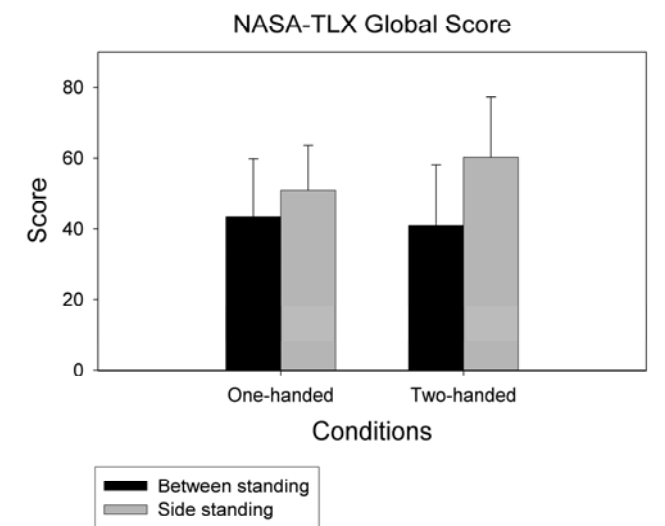| Conditions	Performance time (sec.): Mean (SD)
OHBW	382.4 (46.6)
OHSD	393.4 (33.3)
THBW	298.6 (36.5)
THSD	398.4 (42.9)

Figure 3. Subjective workload analysis using NASA-TLX

LAPAROSCOPIC CHOLECYSTECTOMY POSES PHYSICAL INJURY RISK TO SURGEONS: ANALYSIS OF HAND TECHNIQUE AND STANDING POSITION

Yassar Youssef MD *, Gyusung Lee PhD +, Carlos Godinez MD +, Erica Sutton MD +, Rosemary Klein MA +, Ivan George +, Jacob Seagull PhD +, Adrian Park MD +

From the *Department of Surgery, Sinai Hospital, Baltimore, MD, USA and +Department of Surgery, University of Maryland Medical Center, Baltimore, MD, USA

Introduction: This study compares the effects of surgical techniques (one-handed versus two-handed) and surgeon's standing position (side-standing versus between-standing) during laparoscopic cholecystectomy (LC) and investigates each in regard to surgeons' learning, performance, and ergonomics. There is little homogeneity in how to perform and train for LC. Variations in standing position ("American" or side-standing technique where the surgeon stands on the patient's left versus "French" or between-standing technique where the surgeon stands between the patient's legs) as well as hand technique (one-handed versus two-handed) exist. The two-handed technique refers to the operating surgeon providing exposure of the cystic triangle while using the left hand for manipulation and retraction. The one-handed technique refers to the situation during which the operating surgeon dissects with one hand (generally the dominant hand) and manages the camera/laparoscope with the other. During use of the one-handed technique, the assistant helps to provide exposure and retract the gallbladder and during use of the two-handed technique serves as "camera driver." Our current research augments our previous work which also incorporated assessments of the mental workload exerted during use of both surgical handed techniques and standing positions. That study demonstrated significant association of the side-standing position with high physical demand, effort and frustration and more required effort when the two-handed technique rather than the one-handed technique was used in the side-standing position.

Methods: Thirty-two LC procedures performed by a total of eight subjects on a virtual reality simulator were video recorded and analyzed in this IRB-approved study. All eight subjects were right-handed; five were surgical residents (PGY 2-4); two were minimally invasive surgery fellows; one was an attending. Each subject performed four different procedures so as individual assessment of the following methods was possible: one-handed/side-standing, one-handed/between-standing, two-handed/side-standing, and two-handed/between-standing. Physical ergonomics were evaluated using the Rapid Upper Limb Assessment (RULA) tool. Performance evaluation data generated by both the virtual reality simulator and a subjective survey were also analyzed.

Results: In all 32 procedures performed – regardless of whether the technique used was one- or two-handed - RULA scores were consistently lower (indicating better ergonomics) for the between-standing technique and higher (indicating worse ergonomics) for the side-standing technique. The different scores generated for each anatomical area showed the main disadvantage of the side-standing position to be its detrimental effect on both the upper arms and trunk. The objective, simulator-generated performance metrics demonstrated no differences in either operative time or complication rate among the four methods for performing LC. Survey answers indicated the subjects'

choice to be the two-handed/between-standing method as the best procedural method for teaching and standardization.

Conclusion: Laparoscopic cholecystectomy poses a risk of physical injury to the surgeon. Our research further confirms that the left side-standing position currently in common use in the United States leads to increased physical demand and effort, thus resulting in ergonomically unsound operative conditions for the surgeon. Until further investigations are made, adopting the between-standing position deserves serious consideration as it presents the best short-term ergonomic alternative.

A Ubiquitous Context-Aware Environment for Surgical Training

P. Ordóñez, P. Kodeswaran, V. Korolev, W. Li, O. Walavalkar, B. Elgamil, A. Joshi, T. Finin, Y. Yesha

Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD USA
{ordopa1, palanik1, vkorol1, wenjia1, onkar1, ben8, joshi, finin, yeyesha}@cs.umbc.edu

I. George

Department of General Surgery
University of Maryland Medical Center
Baltimore, MD USA
igeorge@smail.umaryland.edu

Abstract— The age of technology has changed the way that surgeons are being trained. Traditional methodologies for training can include lecturing, shadowing, apprenticing, and developing skills within live clinical situations. Computerized tools which simulate surgical procedures and/or experiences can allow for "virtual" experiences to enhance the traditional training procedures that can dramatically improve upon the older methods. However, such systems do not to adapt to the training context. We describe a ubiquitous computing system that tracks low-level events in the surgical training room (e.g. student locations, lessons completed, learning tasks assigned, and performance metrics) and from these derive the training context. This can be used to create an adaptive training system.

Keywords- context awareness; ubiquitous computing; surgical training.

I. INTRODUCTION

Context aware ubiquitous computing systems must process streams of data arriving from sensors, services, devices and other systems to construct and maintain a model of their environment. If the environment is complex, the volume of data will be large and if the system aspires to be intelligent, the processing over the data may be computationally expensive. In ongoing research, we are designing and implementing a framework for constructing intelligent, context-aware ubiquitous computing systems.

We are pursuing the general technical goals while working with colleagues at the University of Maryland Medical Center (UMMC) to use an evolving project to implement a system named CAST, Context-Aware Surgical Training. CAST is part of the Operating Room of the Future (ORF) [15] project that is housed in the newly opened Maryland Advanced Simulation, Training, Research, and Innovation Center (MASTRI). It is a facility with authentic operating rooms specially renovated/constructed and instrumented to support innovative research and training. We have already constructed and deployed a partial prototype of the CAST system in the MASTRI Center to test the feasibility of our approach, which is described in this paper.

II. THE CAST VISION: BACKGROUND AND MOTIVATION

Traditionally, surgical training has consisted of the resident shadowing senior surgeons and practicing diagnostic and procedural skills on live patients. In 1999, Gorman et al. [4] stated that estimated costs of training chief residents for general surgery alone cost $53 million per year. A long standing debate over the ethics and practicality of such practices is also of concern [2]. Furthermore, statistics like the following demonstrate the need for a dramatic change in clinical pedagogy. A survey of residents and faculty in surgical training programs in 2003 described that more than 87% of the 1,653 responses from residents surveyed indicated that they had an 80 hour work week. 45% reported working more than 100 hours per week. 57% reported that their cognitive abilities had been impaired by fatigue [5]. Furthermore, although apprenticeships have been shown to be very effective, in the case of a surgical procedure, the well being of a patient outweighs the training of the resident.

Computer-enhanced simulations show promise for addressing all of these concerns. However, as Granger [7] states in his dissertation, "The key issue is not whether to creep forward through evolution of digital substitutes, but whether to promote the revolution of clinical practice through the integration of pervasive computing technologies."

Our system aims to improve the training provided by simulators by making them a part of a context-aware training environment. This allows the training process to require less direct intervention from mentors in many of the routine tasks. We aim to reason over sensed data streams to infer context about the events in the training process. In the initial prototype system described in this paper, we focus on laparoscopic surgical training[1]. We track the presence of surgical residents in the training rooms, which training machine they have used and for how long, which lessons they have downloaded etc. This information is then used to guide the students to the practice/tests they need to take. Similarly, recordings of students' hands as they use the laparoscopic trainers as well as the output from the simulators are made available to instructors, who can then view and analyze them, add comments and annotations, and suggest skills on which the

[1] Laparoscopic surgery involves operations in the abdomen that are performed through small incisions rather than the larger ones required by traditional surgical procedures.

trainee needs to focus. Instructor feedback and suggestions can be automatically provided to trainees as podcasts or text messages.

An Electronic Student Record (ESR) provides a centralized repository of student information and progress, and helps infer their appropriate pedagogic context. This record is described in the semantic web language RDF-S, but is presently implemented as a database schema. The ESR will provide a comprehensive summary of the students' progress such as the time spent at each machine, chapters checked out, video captures etc. which will help the instructor to review student performance without physically being present in the training room.

In this system, the student benefits from the guiding elements that can be brought to bear and the real time adaptation of the training. Moreover, the trainer is now able to change their curriculum to meet the needs of their students. For the patient, the movement to disrupt the old and oft-repeated mantra of "See one, do one, teach one" is quite telling. In particular, the steep learning curves of new surgical technologies can now be mastered by trainees outside of live operative settings.

III. RELATED WORK

"William Osler wrote in 'The Principles and Practice of Medicine' in 1982 that: 'To learn medicine without books is to sail an uncharted sea, while to learn medicine only from books, is not to go to sea at all.'" [7]. In the 21st century, the question is can we virtually go to sea?

A. Virtual Reality Training

Parallels have been drawn between pilots and surgeons in that both must be able to respond to potentially life-threatening situations in unpredictable environments [4]. A pilot must be prepared to land a plane when several engines have failed and a surgeon must be able to respond to a cardiac arrest in the middle of open heart surgery. Flight simulators have long been used to train pilots for the worst of circumstances. In fact, the simulators of today are so effective that they are often used to train a pilot on a new version of a plane, and the pilot flies the real plane on a scheduled flight [10].

As a result, surgical simulation is rapidly becoming the standard for surgical training. Training simulations currently exist for endoscopic sinus surgery [3], ossiculoplasty surgery [8], and orthopedic surgery [7] to name a few. Many of these simulations create a virtual reality using video gaming technology. A recent paper in 2007 correlated video gaming skills with the laparoscopic surgical skills [12], although that should not be a reason to relax concerns about the amount of time children spend playing video games.

Some of the aforementioned systems use multimedia and hypermedia to enhance surgical training [7]. Others simply use 2-D video and haptic devices as in most video games [3][8]. Others use a hybrid approach where they combine the 2-D video with a visual awareness of objects and events in a room [11]. Welch et al. are capturing and displaying high-fidelity 3-D graphical reconstructions of the actual time-varying events for the purpose of doing on-line consultation and off-line surgical training [9]. This research could help to provide surgical training and mentoring by specialists to generalist doctors in isolated hospitals in developing countries [6].

The 3-D graphical reconstructions are being stored in Immersive Electronic Books (IEB) for surgical training. Via IEB surgeons can explore previous surgical treatments in 3-D [10]. Thus, in the same way, a pilot can test out a new version of a plane time and time before she flies it, a surgeon can see a surgical procedure and interact with it time and time again until she performs it.

B. Other Training

B-line Medical [1] provides what it describes as a Clinical Skills System that is a comprehensive digital solution for managing and operating a clinical skills center. The system has four major components: user management and content creation, exam management, scoring and reporting, and professional quality audio/video. This system attempts to address the same concerns about efficiency and automation in a surgical training environment. It uses a card swiping mechanism to identify residents and monitor their progress. It is mostly built on existing content management technology, and is not concerned with inferring context from sensed data.

More generally, to the best of our knowledge, none of the existing surgical training systems seek to infer significant events from the sensed data, or to use such data to infer the context of the surgical procedure and create a smart , adaptive surgical training space.
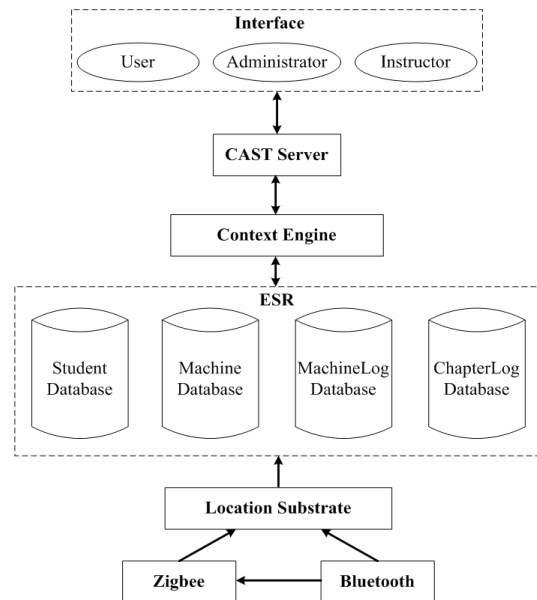
IV. OVERALL ARCHITECTURE

Figure 1.   A high-level overview of the CAST system.

As in any ubiquitous computing system, location plays an important role in CAST. As shown in Figure 1, we use a location substrate consisting of a combination of Zigbee and Bluetooth to provide location information. This location information is then fed into a location database which keeps track of information such as which students were in front of a simulator and for how much time. The simulators in general

require that students complete relevant chapters from the Fundamentals of Laparoscopic Surgery (FLS) training program before working on them. We host the FLS chapters on a web server that students can access through their logins. The chapters checked out are stored as part of the student's Electronic Student Record (ESR), which is updated when students check out chapters through the web interface. When the location substrate detects that a student is standing next to a simulator, it queries the student's context to verify that the student has completed the required FLS chapters. Only if the student has finished the required chapters is he/she allowed to work on the simulator.

Currently, information from sensors, such as training boxes, video recorders, RF tags, and cell phones, provide basic context information. These low level data streams are processed to generate higher-level primitive events, such as a resident entering the training room. A hierarchical knowledge-based event detection system correlates primitive events, resident data, and workflow data to infer high-level events, such as the finishing of a training module. Video streams of the training procedure are time-stamped and labeled with the inferred higher-level events. . These video recordings, location data, and performance data from simulators can be viewed offline by instructors. Moreover some simulator manufacturers are working on providing automated performance evaluation through video metric analysis. This resulting analysis, where available, could also be used as part of the ESR. The resulting ESR will provide trainers of physicians with a permanent record of the training session including an evaluation of the trainee's performance, the duration of the session, the number of times the trainee attempted the module before attempting the exam, and a labeled and time-stamped video of the session.

In a hospital setting with 10-20 residents, the smart training space will monitor the training activities of each resident more closely, and improve the workflow in the training center by allowing residents to sign up for a simulator at an allotted time only if they had the appropriate prerequisite tests/lessons and had been cleared by their mentor. Thus, a trainee surgeon may practice at a simulator during hours of convenience and be evaluated at the end of a session without the need of a trainer.

More generally, an important contribution of our system is that it makes the entire process of surgical training asynchronous. The instructor no longer needs to be physically present with the student during training. Many of the "adaptations" that a physically present instructor would have made (guiding the students to the right simulators and pointing out particular skills they needed to master, for instance) are now done by the system automatically. Moreover, the capture of the data stream from the simulators and the video of the trainee's hands as he/she practice on the laparoscopic simulators allow the evaluation to be done separately as well. This can also help remove the location dependence of surgical training. For instance, a student could do his training at any available simulation center (as long as it is networked with the parent school) and still have his/her procedure reviewed by the instructor back in the parent medical school, or even by an instructor at a different school.

CAST will also alleviate the burden of viewing the entire training video for evaluation even where automatic video metric analysis is not available or possible. It will provide the trainer with a labeled and time-stamped video of each training session that is correlated with the events signaled by the underlying simulator. For instance, the simulator may signal that the cut was made outside the designated area. Since the video timestamps will be correlated to the timestamps of the simulator output, the trainer can jump to portions of the video which are critical.

V. LOCALIZATION

Based on experiments conducted at the MASTRI, the Awarepoint^{TM} system, like most other commercially available location systems such as Ekahau [22], provides room level accuracy which suits the typical requirements of a hospital for asset or personnel tracking. However the CAST system requires finer localization to be able to place a student as being in a position to operate a particular simulator when there may be more than one in a room. We decided to use Bluetooth to provide localization at this granularity. As a result, our system uses a combination of Awarepoint and Bluetooth for localization.

A. Awarepoint^{TM}

Awarepoint seeks to address the limitations in RFID technology and claims to have developed a real-time solution for one of hospitals' major problems, tracking of the movement of their staff, patients, and equipment. Awarepoint's tracking system is based on Zigbee, a high level communication protocol using the IEEE 802.15.4 standard for wireless network [16]. It is designed for radio frequency applications which require a low data rate, long battery life and secure networking. Awarepoint base stations, plugged directly into wall sockets, form a mesh network to deliver data from tags (such as signal strength and identifier) to a server which then uses a proprietary approach to identify the location of the tag. Each trainee is assumed to carry a tag. Awarepoint's standard user interface is a GUI that shows the location of the tags in a facility map. However, this is not appropriate for our purpose. As a part of a collaborative effort with Awarepoint, we have been provided access to their server database and associated SOAP interfaces so that we can directly query the location of a tag.

B. Bluetooth Module

We use Bluetooth to provide machine-level location information so that instructors can query for information such as how much time students spend in front of a machine. We periodically broadcast a Bluetooth device inquiry message, which returns the devices in range which respond to the inquiry. However, this method has high latency and does not necessarily return all Bluetooth devices in range, as some of them may not be listening on the same channel as the inquiry was sent and hence may not respond. As a result we decided to use a different approach. Our approach is motivated by the fact that we are not looking for "any" device but only for devices belonging to trainees. Each trainee is assumed to always have a Bluetooth capable device on him/her and the device address-student association is maintained in the student database. In our method, we periodically initiate connections to a list of MAC addresses obtained from the student database, and if the connection succeeds, we can infer that the corresponding device, and hence student, is in range. This method works well

when the number of students is small, but the time to discover a device in this case grows with the number of students. To reduce the number of MAC addresses to initiate connections to, we use the Zigbee location information which provides room level accuracy, and initiate connections only to devices belonging to trainees currently in the room. When a single trainee is near a simulator, this suffices. However, since multiple trainees could be in the range of a machine, we still need a way to distinguish which one is actually using the machine. We achieve this by displaying a drop down list of students in range and requiring students to log in before using the machine. Thus we provide an additional layer of authentication, when Bluetooth discovery alone is unable to identify a student.

VI. FUNDAMENTALS OF LAPAROSCOPIC SURGERY (FLS)

FLS is "a comprehensive, CD-ROM-based education module that includes a hands-on skills training component and assessment tool designed to teach the physiology, fundamental knowledge, and technical skills required in basic laparoscopic surgery. [14]." It was created by the Society of Gastrointestinal and Endoscopic Surgeons (SAGES) which is accredited by the Accreditation Council for Continuing Medical Education (A.C.C.M.E.) to sponsor Continuing Medical Education for physicians. To the best of our knowledge, FLS is the only CD-based education module that can be used to acquire CME credits. Since our system needs to incorporate the FLS curriculum and move away from its present CD based model, we need to obtain appropriate permissions. While this is being discussed, we have mocked up curriculum to represent the 14 modules in the FLS as shown in Figure 3.

A. Webservice and MySql database

We are hosting the mocked FLS curriculum on an Apache web server. We are using video clips to represent each of the modules. The user has to authenticate herself to the system to check out a training module. We track when students log in, for how long, which chapters they check out, in what sequence, and then log the information in the corresponding tables in the ESR. This information is useful in analyzing student progress. It also allows us to direct students to the tests they need to take and the procedures they need to practice on particular machines. So for instance if a trainee tries to use a simulator for which she has not checked out the appropriate lessons, we prevent her from using it.

VII. IMPLEMENTATION DETAILS

The target development platform for CAST is the Nokia N800 [18]. The N800, pictured in Figure 2, has an impressive set of features such as Bluetooth, WiFi, and an inbuilt camera in a small form factor which makes it an attractive choice for capturing context in the training room. The N800 runs a Debian Linux distribution, which makes developing and porting applications easy. Each simulator has an associated N800 device. It serves as our Bluetooth location base stations. We use the built in camera to capture live video streams of students' training and relay it over WiFi to a central video database for indexing, potential automatic analysis, and review by the instructor. The N800 can also accept the simulator data feeds and stream them to the ESR. Our application code is

mostly written in Python, using python for maemo [17]. N800's built in video chat, RSS feedreader, and podcast applications are also useful in allowing trainee-mentor interactions.

We have implemented Bluetooth localization using the PyBluez [19] module and BlueZ [20] stack on Linux. The Awarepoint server exports location information both through a database and a web service. We currently use the Awarepoint web service to obtain room level location information.

Figure 2. Login page to FLS Interface on Nokia N800.

The ESR is defined using RDF-S. Figure 4 gives a snapshot of the ESR in RDF-S. In the present implementation, we do not use a triple store. Instead, we have a preliminary version of the ESR as a MySQL database to allow for rapid prototyping. As we begin to reason over the sensed data more fully, we will migrate towards a triple store such as Jena [23].
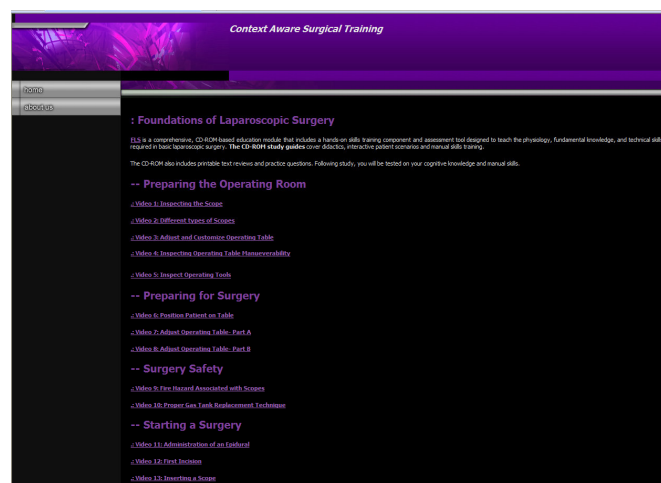

Figure 3. Mock FLS Curriculum

We have implemented and integrated the various modules of the system to create a prototype application. We can populate the student locations from the Awarepoint system, use Bluetooth to recognize trainees, and keep track of (mock) FLS modules checked out. We can also capture the video stream, although we do not yet synchronize it with the data stream captured from the simulators. This is because the simulator data

streams are in proprietary formats for which no public documentation is available. We are presently working with some of the device manufacturers to understand their formats. One of the simulators used in the initial CAST deployment is shown in Figure 5.

```
<rdf:Property rdf:about="&kb;ID"
        a:maxCardinality="1"
        a:minCardinality="1"
        a:range="integer"
        rdfs:label="ID">
        <rdfs:domain rdf:resource="&kb;Student"/>
        <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdfs:Class rdf:about="&kb;PracticalTask"
        rdfs:label="PracticalTask">
        <rdfs:subClassOf rdf:resource="&kb;Task"/>
</rdfs:Class>
<rdfs:Class rdf:about="&kb;Student"
        rdfs:label="Student">
        <rdfs:subClassOf rdf:resource="&a;_system_class"/>
</rdfs:Class>
<rdfs:Class rdf:about="&kb;Task"
        rdfs:label="Task">
        <rdfs:subClassOf rdf:resource="&a;_system_class"/>
</rdfs:Class>
<rdfs:Class rdf:about="&kb;TheoreticalTask"
        rdfs:label="TheoreticalTask">
        <rdfs:subClassOf rdf:resource="&kb;Task"/>
</rdfs:Class>
<rdf:Property rdf:about="&kb;performed"
        a:range="cls"
        rdfs:label="performed">
        <rdfs:domain rdf:resource="&kb;Student"/>
        <a:values rdf:resource="&kb;Task"/>
        <rdfs:range rdf:resource="&rdfs;Class"/>
</rdf:Property>
<rdf:Property rdf:about="&kb;precedes"
        a:range="cls"
        rdfs:label="precedes">
        <a:values rdf:resource="&kb;Task"/>
        <rdfs:domain rdf:resource="&kb;Task"/>
        <rdfs:range rdf:resource="&rdfs;Class"/>
</rdf:Property>
```

Figure 4. Snapshot of ESR in RDF-S.

Figure 5. A Laparoscopic Training Simulator with an N800.

VIII. FUTURE WORK

In the near future, we will evaluate the impact of such a context aware, ubiquitous system for surgical training in collaboration with colleagues at the University of Maryland Medical System. We will consider the security, effectiveness and efficiency of the system. Other simulators at MASTRI include a ProMIS[TM] surgical simulator as well as a METI[TM] Human Patient Simulator[TM]. The ProMIS[TM] surgical simulator includes performance metrics that have been validated for use with the SAGES FLS program that we intend to use in the ESR. The METI[TM] Human Patient Simulator[TM] provides a log of vital signs during a surgical training procedure. We aim to expand CAST to capture data from these sources as well.

REFERENCES

[1] "Clinical Skills System", B-Line Medical: Comprehensive Digital Solutions, http://www.blinemedical.com/solution/clinicalskills.

[2] M.A Reznek, P. Harter, and T. Krummel, "Virtual reality and simulation: training the future emergency physician," Acad. Emer. Med., vol. 9, issue 1, pp. 78-87.

[3] M. . Fried, R. Satava, S. Weghorst, A.G. Gallagher, C. Sasaki, D. Ross, M. Sinanan, J.I. Uribe, M. Zeltsan, H. Arora and H. Cuellar, "Identifying and reducing errors with surgical simulation," Qual. Saf. Health Care, vol. 13, pp. 19-26, 2004.

[4] P.J. Gorman, A.H. Meier, T.M. Krummel, "Simulation and virtual reality in surgical education: real or unreal?," Arch. Surg., vol. 134, pp. 1203-1208, Nov 1999.

[5] M.J. Niederee, J.L. Knudtson, M.C. Byrnes, S.D. Helmer, R. S. Smith, "A survey of residents and faculty regarding work hour limitations in surgical training programs,"Qual. Saf. Health Care, vol. 13, pp. 19-26.

[6] F. H Garlick, "Surgical training of doctors in their own isolated hospital," Aust. N.Z. J. Surg., vol. 70, pp. 456-458, 2000.

[7] S. Grange, "A virtual university infrastructure for orthopaedic surgical training with integrated simulation," PhD, Engineering, Mathematics and Computing, University of Exeter, United Kingdom, 2006.

[8] K. Kalicki, F. Starzynski, A. Jenerowicz, K. Marasek, "Simple ossiculoplasty surgery simulation using haptic device," *mue*, pp. 932-936, 2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07), 2007.

[9] G. Welch, R. Yang, B. Cairns, H. Towles, A. State, A. Ilie, S. Becker, D. Russo, J. Funaro, D. Sonnenwald, K. Mayer-Patel, B. D. Allen, H. Yang, E. Freid, A. van Dam, and H. Fuchs, "3D Telepresence for Off-Line Surgical Training and On-Line Remote Consultation," Proceedings of ICAT CREST Symposium on Telecommunication, Teleimmersion, and Telexistence, The University of Tokyo, Tokyo, Japan, December 2004. Invited submission.

[10] G. Welch, R. Yang, S. Becker, A. Ilie, D. Russo, J. Funaro, A. State, K.Low, A. Lastra, H. Towles, B. Cairns, H. Fuchs, and A. van Dam. "Immersive Electronic Books for Surgical Training, " IEEE Multimedia, vol. 12, no. 3, pp. 22–35, July–September 2005.

[11] A. Ilie, k. Low, G. Welch, A. Lastra, H. Fuchs and B. Cairns, "Combining Head-Mounted and Projector-Based Displays for Surgical Training," Presence: Teleoperators and Virtual Environments, vol. 13, no.2, pp. 128-145, April 2004.

[12] J.C. Rosser, Jr., P.J. Lynch, L. Cuddihy, D.A. Gentile, J. Klonsky, R. Merrell, "The impact of video games on training surgeons in the 21st century," Arch. Surg., vol. 142, no. 2, pp. 181-186, Feb 2007.

[13] "The Awarepoint™ Solution", Awarepiont: Real-time Awareness Solutions, http://awarepoint.com/Healthcare/Solution.html.

[14] "What is FLS", Fundamentals of Laparoscopic Surgery, http://www.flsprogram.org/.

[15] "OR of the Future", University of Maryland Medical Center, http://www.umm.edu/center/or_of_future.htm.

[16] "Zigbee", Wikipedia, http://en.wikipedia.org/wiki/Zigbee.

[17] "Maemo", http://maemo.org/

[18] "Nokia N800", http://www.nseries.com/n800/.

[19] "PyBluez", http://org.csail.mit.edu/pybluez

[20] "BlueZ', http://www.bluez.org/

[21] "RDF-S", http://www.w3.org/TR/rdf-schema/

[22] "ekahau", http://ekahau.com/

[23] "Jena-A Semantic Web Framework for Java", http://jena.sourceforge.net

[24] "ProMIS™ surgical simulator", http://www.haptica.com/id11.htm

[25] "Human Patient Simulator", METI Medical Education Technologies, Inc., http://www.meti.com/Product_HPS.html

Video Summarization of Laparoscopic Cholecystectomies

Michael A. Grasso, MD, PhD, Timothy Finin, PhD, Anupam Joshi, PhD, Yelena Yesha, PhD

Department of Computer Science, University of Maryland Baltimore County, Baltimore, Maryland University of Maryland School of Medicine, Baltimore, Maryland

ABSTRACT

We compared image features with a distance metric to identify the critical view of a laparoscopic cholecystectomy. Our initial results were promising, but more work needs to be done to increase accuracy. We are currently experimenting with particle analysis, edge analysis, and support vector machines as ways to create a more robust image classifier.

BACKGROUND

Laparoscopic surgery is a minimally invasive technique that is the method of choice for a number of surgical procedures. Patients who undergo laparoscopic surgery have smaller scars, reduced pain, and a quicker recovery. The laparoscopic approach, however, is more technical challenging and has more demanding training requirements [1]. Our overall goal is to develop a software tool to assist with video-based assessment of surgical trainees. We present an initial feasibility study, where we compared image features with a distance metric to identify the critical view during a laparoscopic cholecystectomy (surgical procedure to remove the gall bladder) [2]. The critical view is an important validation step in the surgery when the essential anatomy has been identified. Related efforts include the segmentation of hysteroscopy video [3] and echocardiogram video [4].

METHODS

We randomly selected 378 representative images from 5 laparoscopic cholecystectomy videos, with 104 of the images collected near the critical view. We analyzed 49 separate spectral and textural features. A surgeon reviewed the videos to identify the critical view images. We used FFmpeg and ImageJ to extract images and features. We applied the Jeffrey Divergence to the data 5 times, each time using a critical view image from a different case as our basis for comparison. We chose threshold values empirically to maximize sensitivity and specificity.

RESULTS

A summary of the best image features is shown in Table 1. They include color histogram (color distribution), energy (pixel uniformity), entropy (pixel complexity), contrast (local variation), and correlation (linear patterns).

Image Feature Sensitivity Specificity
Color Histogram 71.5% 72.2%; Energy 66.4% 67.2%; Entropy 60.9% 62.5%; Contrast 62.4% 62.6%; Correlation 62.2% 60.0%; Table 1. Sensitivity and specificity of image features.

DISCUSSION

Our initial results show promise with a sensitivity and specificity up to 72%. Color histograms and textural energy performed the best. Accuracy, however, must improve

before there can be any practical application of this approach. When interpreting our results it is important to consider several limitations. The study was small in size. The cases were restricted to a single academic medical center. Finally, our comparisons were limited to one feature from one image at a time. We are currently working on a more robust image classifier using support vector machines, and extracting more robust features through particle analysis and edge analysis.

REFERENCES
1. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. Br J Surg. 2004 Dec;91(12):1549-58.
2. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Transaction on Systems, Man, and Cybernetics. 1973 Nov; 3(6):610-621.
3. Scharcanski JN, Neto WG, Cunha-Filho JS. Diagnostic hysteroscopy video summarization and browsing. Proceedings 27th IEEE-EMBS. 2005;:5680-5683.
4. Roy A, Sural S, Mukherjee J, Majumdar AK State-based modeling and object extraction from echocardiogram video. IEEE Transaction on Information Technology in Biomedicine. 2008 May;12(3):366-376.

Cognitive Simulation in Virtual Patients

Sergei Nirenburg, Marjorie McShane, Stephen Beale

University of Maryland, Baltimore County
{sergei,marge,sbeale,tomohara}@umbc.edu

Bruce Jarrell, George Fantry

University of Maryland School of Medicine
bjarrell@som.umaryland.edu
gfantry@medicine.umaryland.edu

The Maryland Virtual Patient (MVP) project aims to create an environment where a physician user can manage a complex virtual patient who is suffering from one or more diseases. Once developed, this environment will allow for multiple capabilities, including trial-and-error learning, tutored learning, and assistance with problem solving in treating real patients. Last year we presented our progress by demonstrating virtual patients suffering from complex esophageal diseases that progressed over time. These patients behaved in a clinically appropriate fashion and reacted in realistic ways to interventions. The physician could use drop-down menus to select queries, tests and interventions, as well as observe the responses and subsequent disease progression. Simulation was supported by an intelligent agent in the MVP whose main component is a model of normal and abnormal physiology. During the past year, we have developed three additional components of the intelligent agent: a) two types of perception: perception of stimuli originating inside of the body (interoception) and the perception of natural language communication (language perception), b) a model of cognitive decision-making and c) a model of verbal and simulated physical action.

Perception: The physiological component of the agent communicates with the cognitive component as simulated interoception to produce symptom perception by the cognitive agent. Language perception allows the cognitive agent to understand the meaning of inputs it obtains from the physician user.

Cognitive Decision-Making: This module of the system uses several types of input and knowledge to model agent decision making, including: interoception; input from the physician; the resident knowledge possessed by the agent; and the agent's personality traits.

Simulated Action: In the current simulation, the action can be verbal (for example, a response to the physician's question), or simulated physical (for example, taking medication or presenting at the physician's office). Our objective is to model these actions in a way that would be natural for people.

The variables used in building this model of an intelligent virtual patient include:
- Life goals, such as the desire to be healthy
- Character traits, such as
  - Attitude toward visits to the physician
  - Courage to be treated
  - Trust in the physician's skill
  - Suggestibility with respect to the physician's recommendations
- Physiological traits, such as

- o   Tolerance to pain
- o   Tolerance to symptoms
- o   Tolerance to external stressors
- Intellectual traits, such as
  - o   Memory/forgetfulness of symptoms and events
  - o   Knowledge about the disease
  - o   Knowledge about the tests and interventions
  - o   Retention of knowledge gleaned in conversation with the physician.

Language Processing Capabilities.

When free-text input is received from the physician, it is interpreted and assigned a formal text meaning representation. During this process, both the meaning and intention of the input are determined. Indirect speech acts, such as implied questions not stated in a question format, are handled appropriately. Physician input that is currently interpreted by the MVP agent includes:

1. a request for physiological data (test results), perception of symptoms, and memory of health events
2. a request for permission to carry out a test or intervention
3. a request to perform an intervention
4. a request to return to see the physician at a later date
5. a response to an MVP question
6. unsolicited knowledge provided to the MVP by the physician

Natural language output is provided to the physician after processing by the MVP agent. All responses are based upon evaluations of the physiological state of the MVP in concert with the perception of symptoms and cognitive functions. Potential MVP agent output includes:

1. providing requested data
2. requesting additional information
3. inquiring about other treatment options available at a given time
4. agreeing to or refusing to submit to the physician's suggestion for a test or intervention
5. storing new knowledge in the MVP memory
6. presenting to the physician in response to an intolerable state of health (as defined by the MVP) for a first or subsequent visit
7. presenting to the physician at a later date in response to the physician's request

We will demonstrate this process by communicating in natural language with two simulated patients. The first patient is a knowledgeable individual who: a) foresees additional information the physician might desire, such as the frequency of a symptom when asked if he experiences that symptom, and b) desires a lot of additional information about what the user proposes to do. In addition, the first patient learns from the encounter, so that the next time the physician suggests an intervention which the patient already knows about, the patient agrees without further questioning because he remembers the results of the original decision-making process. The second patient: a) is a trusting individual who essentially agrees to all suggestions with no

questions asked and b) responds to all questions quite literally, not providing any additional explanatory information.

In summary, we will display several accomplishments:

1. The MVP retains its previous physiological complexity while gaining the ability to communicate in natural language and to incorporate interoception, cognitive traits and behavioral traits in its own health care decision-making
2. The MVP has personality traits that give it curiosity, free-will and other human characteristics when responding to the physician
3. The MVP can learn from explanations and actions by the physician and use that knowledge in future decisions about health care
4. The MVP interaction involves a conversation that is becoming very realistic.

LIVE AUGMENTED REALITY– A NEW VISUALIZATION METHOD FOR LAPAROSCOPIC SURGERY USING CONTINUOUS VOLUMETRIC CT

Raj Shekhar, PhD,[1,2,3] Venkatesh Bhat, MS,[1,2] Omkar Dandekar, PhD,[1,2] Mathew Philip, MS,[1,2] Peng Lei, MS,[1,3] Carlos Godinez, MD,[4] Erica Sutton, MD,[4] Ivan George, MS,[4] Steven Kavic, MD,[4] Reuben Mezrich, MD, PhD,[1] Adrian Park, PhD[4]

[1]Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, Baltimore
[2]Department of Electrical and Computer Engineering, University of Maryland, College Park
[3]Fischell Department of Bioengineering, University of Maryland, College Park
[4]Department of Surgery, University of Maryland School of Medicine, Baltimore

Corresponding Author: Raj Shekhar, PhD
                           Department of Diagnostic Radiology and Nuclear Medicine
                           University of Maryland School of Medicine
                           22 S Greene Street
                           Baltimore, MD 21201

                           rshekhar@umm.edu
                           (410) 706-8714 phone
                           (410) 706-8724 fax

Abstract

Background
Current laparoscopic images are rich in surface detail but lack information on deeper structures invaluable to the operating surgeon. We have presented here a novel method to highlight these structures during laparoscopic surgery using continuous multislice computed tomography (CT). This has resulted in a more accurate augmented reality (AR) approach, called Live AR, that merges three-dimensional (3D) anatomy from live low-dose intraoperative CT with live images from the laparoscope.

Methods
We conducted a series of swine procedures in a CT room with a fully equipped laparoscopic surgical suite. A 64-slice CT scanner with continuous scanning capability helped image the surgical field approximately once per second. The procedures began with a contrast-enhanced, diagnostic-quality CT scan (initial CT) of the liver followed by continuous intraoperative CT and laparoscopic imaging with an optically tracked laparoscope. Intraoperative anatomic changes included user-applied deformations and those from breathing. Through deformable image registration, an intermediate image processing step, we warped initial CT to spatially align with the low-dose intraoperative CT scans. Registered initial CT was then rendered and merged with laparoscopic images to create Live AR.

Results
Superior compensation of soft-tissue deformations in our methodology led to more accurate spatial registration between laparoscopic and rendered CT images in Live AR than in conventional AR. Furthermore, substitution of low-dose CT with registered initial CT helped visualize the vasculature continuously and offered the potential of at least 8-fold reduction in intraoperative x-ray dose.

Conclusions
We proposed and developed Live AR, a new surgical visualization approach that merged rich surface detail from a laparoscope with instantaneous 3D anatomy from continuous CT scanning of the surgical field. Through innovative use of deformable image registration we also demonstrated the feasibility of continuous visualization of the vasculature and considerable x-ray dose reduction. This study provides motivation for further investigation and development of Live AR.

Keywords: laparoscopic surgery, augmented reality, surgical visualization, continuous CT, image registration, x-ray dose reduction

Introduction

Minimally invasive laparoscopic surgeries present an attractive alternative to conventional open surgeries, and, have been shown to lead to improved outcomes, less scarring and significantly faster patient recovery [1, 2]. For certain surgical procedures, such as cholecystectomy, they have become the standard of care [3]. Despite their success and increasing application to treat various pathological conditions, the visualization of the surgical field in some regards is more challenging in laparoscopic surgeries than in open surgeries [4, 5]. Current laparoscopic images are rich in surface detail but provide no information on deeper features. A surgeon is thus unable to see inside or around exposed surfaces, potentially affecting the precision of current-generation laparoscopic surgeries. Intraoperative appreciation of visible anatomy along with awareness of underlying structures and vasculature would be invaluable to the operating surgeon [5]. The reduced tactile feedback and limited visual displays of minimally invasive surgeries has only heightened the need for improved visualization of target anatomy and visually imperceptible adjacent structures. Laparoscopes are fundamentally limited in providing this information.

To perform true 3D visualization, a volumetric image of the surgical field is essential—the type of data which is basic to modern computed tomography (CT) and magnetic resonance (MR) imaging, but not to laparoscopes. Prior attempts have utilized CT and MR imaging data sets of the relevant anatomy to introduce three-dimensional (3D) visualization to minimally invasive surgeries [6-8], but, because CT and MR imaging scanners are generally unavailable in an operating room and during surgery, these studies have used preoperative CT/MR imaging data sets. Brilliant 3D renderings from these data sets can be and have been generated. Furthermore, steps have been taken to superimpose these renderings on laparoscopic video to create augmented reality (AR), which provides a larger context to small field-of-view of laparoscopy and helps visualize the underlying vessels and other structures.

These studies have taken care to bring preoperative CT/MR data sets into alignment with the patient and the laparoscope's frame of reference. However, a problem with the preoperative images is that they are not reflective of the ever-changing surgical field. Guiding surgeries and basing critical surgical decisions on the 3D rendering from an old snapshot of the target anatomy therefore may be inaccurate and unsafe. Moreover, this problem will persist as long as preoperative CT/MR imaging continues to be used as a proxy for the dynamic surgical field.

The correct approach to solving this problem is to render live, real-time 3D images of the surgical field—an approach we have proposed and whose feasibility we have tested in this study. We use a 64-slice CT scanner with continuous scanning capability for intraoperative imaging. Intraoperative visualization during laparoscopy is improved through AR that uses 3D renderings of the anatomy scanned with live, intraoperative CT, a capability we call *Live AR*. Superimposition of such 3D views based on instantaneously acquired CT on the laparoscopic video after accounting for proper alignment has the potential to reveal hidden structures accurately with their latest location. Although computationally and practically more challenging, the Live AR visualization does not suffer from the limitations of previously reported AR efforts. With the advent of multislice CT scanners, continuous volumetric CT at high frame rates is becoming possible. The continual trend toward more slices (i.e., greater volumetric coverage per

rotation) and higher frame rate (i.e., greater temporal resolution) will make CT even more suitable for this surgical imaging task.

We have presented here an offline feasibility testing of Live AR possible with continuous intraoperative CT. A concern with the use of continuous CT is patients and surgeons receiving excessive levels of radiation exposure. We also describe a strategy to reduce the radiation dose based on registration of initial and intraoperative CT scans. Our results suggest that radiation dose can be reduced to clinically acceptable safe levels and, with further technical development, Live AR can be implemented for routine clinical use. We conclude this article with a discussion of our results, strengths of our proposed strategy, and future directions of our research.

Materials and Methods

A team of surgeons, engineers, radiologists, and supporting staff collaborated to develop and demonstrate the proposed Live AR visualization concept for laparoscopic surgery in the swine. The animal protocol was approved by the institutional animal care and use committee and the experiments were conducted under the vigilance of the veterinary staff. The experiments were performed in a CT room with a 64-slice CT scanner (Brilliance 64, Philips Healthcare, Cleveland, OH). A fully equipped laparoscopic surgery suite with necessary instruments and surgical tools was assembled in the CT room before each experiment. Fig. 1 shows a picture of the typical experimental setup, details of which are described below.
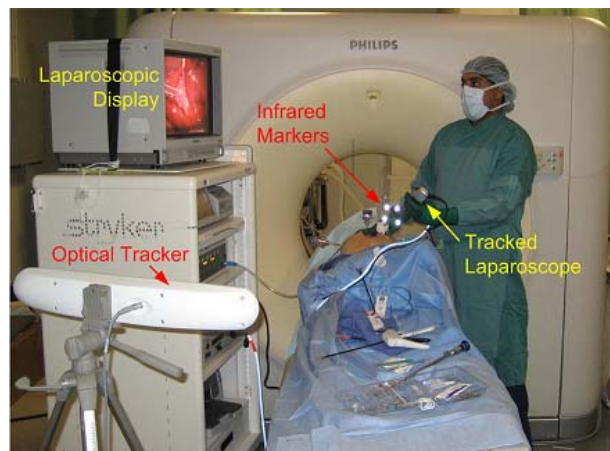

Fig. 1 Typical experimental setup for continuous CT-based laparoscopic surgery with Live AR visualization. The major equipments include a CT scanner, laparoscopic imaging system, and an optical tracker for tracking the laparoscope in CT coordinates.

Imaging Protocol

Fig. 2 shows our proposed imaging protocol that includes a dose reduction strategy. After the animal has been anesthetized and prepared (preparation includes insufflation) for laparoscopy, we acquire a contrast-enhanced CT scan of the liver at the standard diagnostic dose (x-ray tube voltage of 120 kV; x-ray tube current of 200-250 mA, depending on the animal's weight). The

use of the contrast agent ensures that the desired hepatic vessels are highlighted in the CT scan. We adjusted the delay to maximize arterial phase enhancement. We have termed this contrast-enhanced, diagnostic-quality CT scan the *initial* CT scan.
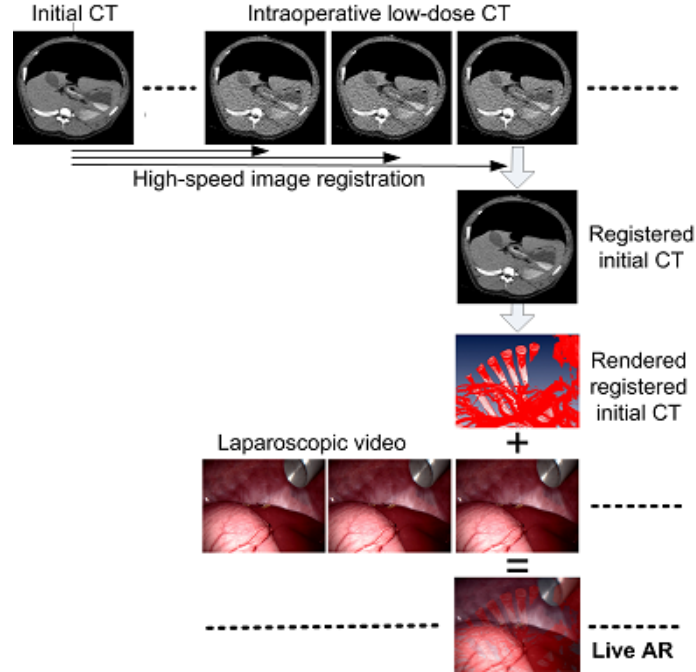
Fig. 2 A flow diagram of our proposed imaging protocol and the built-in dose reduction strategy.

After surgery begins, we propose CT scanning of the surgical field continuously and repeatedly, except the CT scanner is now operated at a much lower dose. We refer to these subsequent low-dose CT scans as intraoperative CT scans. The intraoperative CT scans are not contrast-enhanced because the contrast agents are short-acting and cannot be administered repeatedly without causing stress and harm to the kidneys and other critical organs.

The next step in our protocol is to register, or spatially align, initial and intraoperative CT scans rapidly, which allows us to warp the initial CT scan such that it matches the instantaneous intraoperative anatomy. The registered initial CT scan, which has clinically acceptable image quality and contains the vasculature information, is then substituted for the intraoperative CT scan. This scan is subsequently rendered and superimposed on the corresponding laparoscopic image, accounting for correct camera orientation and optics. By repeating this process for each intraoperative CT scan, the protocol leads to accurate and up-to-date AR visualization throughout the surgery or Live AR.

The dose reduction results from the proposed use of low-dose CT during the surgery. In this study we have experimented with three different dose or x-ray tube current settings to determine the lowest acceptable dose setting.

Deformable Image Registration and Validation

The registration of initial CT and intraoperative CT was performed using a fully automatic algorithm we have reported previously [9]. The algorithm operated in the deformable mode to account for soft-tissue deformations expected in the abdomen. For efficiency and eventual clinical implementation standpoints, we used a previously reported high-speed implementation of this algorithm [10]. Prior to image registration, low-dose CT scans were preprocessed using an anisotropic diffusion filter to reduce noise [11]. The initial alignment before performing image registration was determined by the location data saved with the CT images.

The quality of image registration was judged visually by comparing fused initial CT and intraoperative CT scans before and after registration. To enable objective validation of image registration, we implanted four-to-six point fiducials (2-3 mm pieces of a nonmetallic guidewire) randomly into the liver parenchyma under ultrasound guidance and sutured two small calcium markers on the surface of the liver. The average distance between homologous markers, before registration, was a measure of initial misregistration. After image registration, the same average distance, called target registration error (TRE), determined the accuracy of image registration. Both initial misregistration and TRE were computed.

Procedure for Creating and Validating Augmented Reality

AR is the overlay of optical image from the laparoscope with a computer-generated image of the CT scan. The location and orientation of the laparoscope and the optics of its built-in camera determine the appearance of the laparoscopic image. For accurate spatial registration between the two types of images in AR, the CT scan must be rendered using a virtual camera that mimics the optics of the actual camera and is also be placed at exactly the same location and in the same orientation as the actual camera. We achieved this by optical spatial tracking of the laparoscope for its 3D location and orientation and standard camera calibration for determining the camera optics.

The first step in AR visualization is to crosslink the CT coordinate system and the coordinate system of the laparoscope. Once initialized, the coordinate system of the CT scanner remains fixed. Because the laparoscope is manually operated, its coordinate system is movement dependent and variable. We followed the freehand movement of the laparoscope with an optical tracker (Polaris Spectra, Northern Digital, Waterloo, Canada) that was mounted on a tripod and was kept stationary in the CT room throughout a given experiment. The rigid structure of the laparoscope allowed tracking it by attaching infrared markers placed on its length external to the animal's body. Fig. 1 shows the placement of manufacturer-provided markers on the laparoscope. The optical tracking system was able to track the laparoscope as long as the line of sight between the optical tracker and the markers on the laparoscope was maintained. A PC (termed control PC) controlled the optical tracker and was also fitted with a video frame grabber to capture and digitize the laparoscopic video. The control PC acquired synchronized spatial tracking data of the laparoscope and digitized video frames produced by it.

The determination of camera optics followed standard steps [12, 13]. Specifically, we used an open source camera calibration toolbox [14] that generated camera parameters that permitted the

rendering software (Amira, Visage Imaging, San Diego, CA) to generate CT views. The distortion parameters from camera calibration helped undo the peripheral distortion in the laparoscopic images before superimposing them on rendered CT images.

The registration of laparoscopic and CT images was achieved through first principles, i.e., matching camera optics and location and orientation. To visually verify this registration, we reused the two aforementioned small Calcium markers sutured to the surface of the liver. In our experiments, we took steps to ensure that these markers were visible by the laparoscope as well as contained in the CT field of view. The spatial overlapping of these markers in the AR views constituted an independent verification of our methodology.

Experimental Details

We conducted six animal experiments. The experiments were incremental in nature in that successive experiments helped develop, test, and refine our methodologies. The goal in our experiments was to collect all the necessary data for testing and validating various methodological steps, determining the lowest acceptable dose setting, and creating examples of Live AR visualization, all in an offline fashion. To differentiate Live AR from conventional AR and to demonstrate the former's better accuracy, we created examples of both.

After initial equipment setup and animal preparation, the major steps in our experiments, in order, were (1) calibration of the coordinate systems of the CT scanner and the optical tracker, (2) implantation of wire markers in the liver parenchyma, (3) insufflation, (4) implantation of Calcium markers on the liver surface, (5) acquisition of initial contrast CT, and (6) acquisition of intraoperative CT scans.

For Live AR, the intraoperative CT imaging included acquisition of 100 consecutive volumes (stacks of 64 slices with a 4-cm longitudinal coverage) separated by 1.1 s. The intraoperative imaging was repeated for three x-ray tube current settings. Accompanying intraoperative imaging was also continuous laparoscopic imaging from a fixed location. For creating conventional AR, intraoperative acquisitions were single (not continuous) snapshots of the anatomy. Immediately after the CT, the tracked laparoscope was continuously moved around the anatomy of interest and the resulting video, lasting approximately 1-3 min, was recorded. These steps were repeated, as before, for three dose settings.

The duration between initial CT and intraoperative CT varied between 10 minutes to 2 hours. After initial CT the liver as a whole was manipulated to simulate anatomic shifts from the time of initial CT. For creation of Live AR, we overventilated the animal to cause additional breathing-induced anatomic differences. Live AR, in principle, is capable of following the breathing-induced liver motion that is observed in both continuous CT and laparoscopic imaging. Because the CT scan used for conventional AR was a snapshot, conventional AR showed misregistration arising from breathing phase differences. The spatial overlapping of the two surface markers helped compare the two approaches. For each animal, our experiments allowed creation of three Live AR and three conventional AR animations, one for each of the three dose settings. Static frames as well as animated segments are presented next.

Results

Accuracy of Deformable Image Registration

Deformable image registration plays a crucial role in both reducing the intraoperative radiation dose and enhancing the vessels intraoperatively for creating Live AR. The registration must be accurate for any error would impact the accuracy of Live AR and consequently the accuracy with which structures can be targeted during surgery. Fig. 3 shows the accuracy of image registration qualitatively. The first column shows an axial slice of the initial diagnostic-quality and contrast-enhanced CT scan. This slice is the same for all rows, which show the results of deformable image registration of the initial CT with intraoperative CT acquired at three different doses: 200 mA (high dose), 75 mA (intermediate dose), and 25 mA (low dose). The second column shows the axial intraoperative CT slice from the same longitudinal location as that of the initial CT slice in the first column. The third column shows checkerboard fusion of the initial CT and intraoperative CT before registration. The discontinuities at tile boundaries indicate that, despite the same longitudinal location, the images are misaligned because of intervening liver motion and deformation from user manipulations and breathing. After image registration, however, the misregistration disappears as is apparent in the fusion images in the fourth column. A second important finding here is that no visually noticeable difference is present in the quality of image registration when the dose is varied from high to low, indicating that the quality of image registration is independent of the intraoperative dose setting for the range explored and that 25-mA CT can be used intraoperatively as effectively as higher-dose CT.
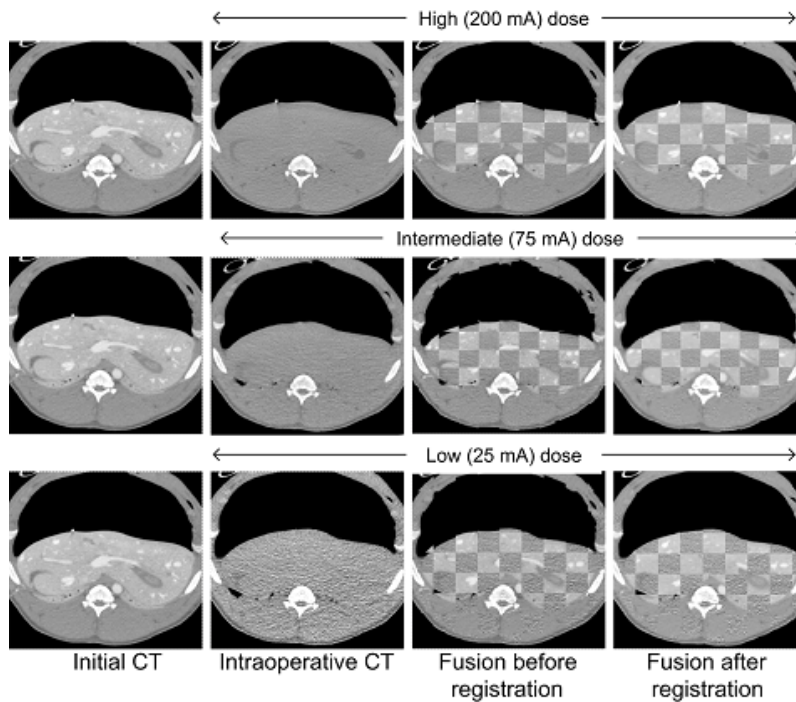
Fig. 3 Registration of initial CT with intraoperative CT acquired at three doses (top row = 200 mA; center row = 75 mA; bottom row = 25 mA). The fusion images before and after registration suggest that image registration performed acceptably at all doses.

Table 1. Initial misregistration and target registration error after deformable image registration.

| Intraoperative CT Dose (mA)	Initial Misregistration (mm)	Target Registration Error (mm)
200	3.12	1.47
75	3.63	1.67
25	3.25	1.45

The accuracy of image registration was quantitatively examined with the aid of the implanted markers. For each intraoperative dose, image registration reduced initial misregistration of greater than 3 mm to an acceptable level of approximately 1.5 mm (see Table 1). Furthermore, the post-registration TRE is relatively independent of the intraoperative dose, indicating again the feasibility of performing intraoperative CT at the lowest dose setting of 25 mA. The procedures described here demonstrate the feasibility of substituting low-dose intraoperative CT with modified initial CT. The modified initial CT scan, when rendered, permits 3D visualization of the hepatic vasculature, which is shown next.

Live AR versus Conventional AR

We have created examples of both Live AR and conventional AR to draw distinctions between the two and to demonstrate the former's better accuracy. We start with conventional AR. In our implementation of conventional AR, we used a single low-dose CT scan of the intraoperative anatomy acquired immediately before a period of AR visualization during which the laparoscope was manipulated and moved around inside the retroperitoneal cavity. As discussed, the low-dose CT scan was substituted with registered initial CT scan. The animal breathed normally during the CT acquisition and the ensuring period of AR visualization.

Fig. 4 displays the laparoscopic, CT and AR views of the liver and the surrounding anatomy. Top and bottom rows show the same sequence of views but for two different time instants (i.e., laparoscope positions and orientations). The liver surface in the CT rendering was made transparent to emphasize the vasculature. Note that the hepatic vessels (as well as the ribs) under the liver surface, invisible in the laparoscopic image, are visible in the CT view. It is also important to note the benefit of image registration for the visualization of the vasculature. The vessels are not enhanced in actual intraoperative CT. Image registration allows using the initial CT for 3D visualization of the intraoperative anatomy while also retaining the vasculature information through contrast enhancement. The AR visualization preserves the surface texture information and optical depth cues from the laparoscope while also exposing the vasculature.

Fig. 5 shows two static frames corresponding to two different time points during the period of Live AR visualization. The CT is rendered, as before, by making the liver surface transparent. Important here is to note that during Live AR, two component views remain spatially aligned. Because the CT scanner we used could scan only a 4-cm thick section of the abdomen continuously, fewer underlying vessels and ribs were exposed during Live AR compared to the conventional AR example above. Small field of view notwithstanding, continuous CT scanning allowed following the exaggerated breathing motion making Live AR more accurate as discussed next. An animation of Live AR can be viewed by clicking on this link.
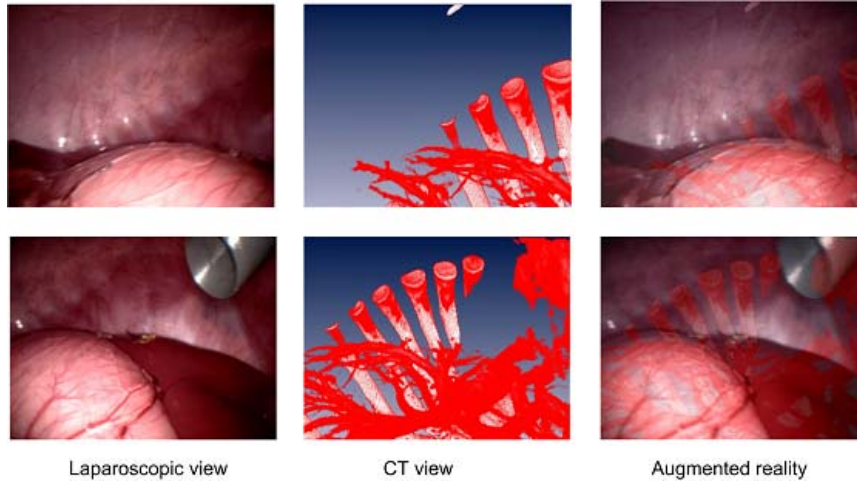
Fig. 4 Laparoscopic (left column), CT-generated (middle column), and AR (right column) views for two different time instants during conventional AR. The AR views combine the strengths of the two visualization techniques.
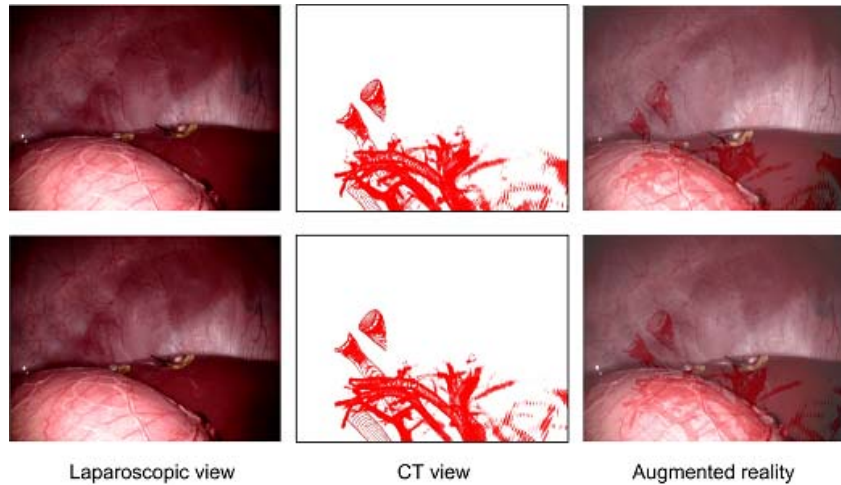

Fig. 5 Laparoscopic (left column), CT-generated (middle column), and AR (right column) views for two different time instants during a Live AR episode. The CT view is capable of revealing the underlying vasculature, visualization of which is beneficial to laparoscopic surgeons. The AR views combine the strengths of the two visualization techniques.

Improved Accuracy of Live AR

Our experimental results confirm the expected improved accuracy of Live AR compared with conventional AR. The registration of the laparoscopic and CT views was achieved through first principles as described in the methods section. However, because CT is not repeated during the length of conventional AR, the spatial registration between the laparoscopic view and the rendered CT view was not perfect. This is evident from the large misregistration of a surface marker seen in the AR view in Fig. 6. The views from the two modalities do not overlap perfectly. Furthermore, the degree of this misregistration is variable (compare top and bottom rows) and, in fact, dependent on the phase of breathing. It is less pronounced when the breathing phase in which the laparoscopic image was acquired is close to the phase in which the intraoperative CT was acquired. The misregistration is accentuated when the two phases differ.

Live AR addresses this misregistration problem inherent in conventional AR because CT is continuously acquired and the temporal separation between the CT scan used to create AR and the corresponding laparoscopic frame is minimized. As before initial CT scan was registered with each incoming frame of continuous CT and rendered to visualize the intraoperative anatomy. Marker-based verification, shown in Fig. 7, confirms superior registration of component images in Live AR. There surface markers were indeed not used to align the two individuals views, rather used merely for verification.

Laparoscopic view   CT view   Augmented reality
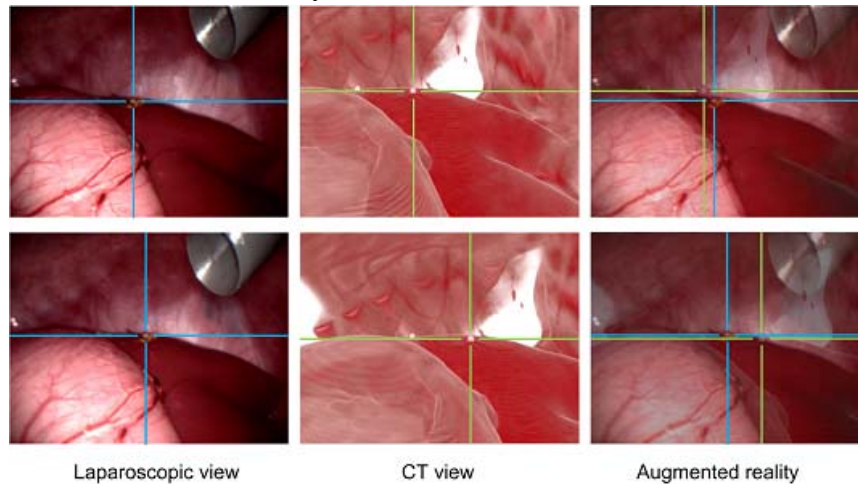
Fig. 6 Conventional AR views (right column) from two time instants shown in top and bottom rows. The two crosshairs pointing to a surface marker reveal large misregistration which is caused by breathing that the conventional AR technique is incapable of correcting. A comparison of results in top and bottom rows shows that the degree of misregistration is variable and confirms breathing as its source.

Laparoscopic view   CT view   Augmented reality
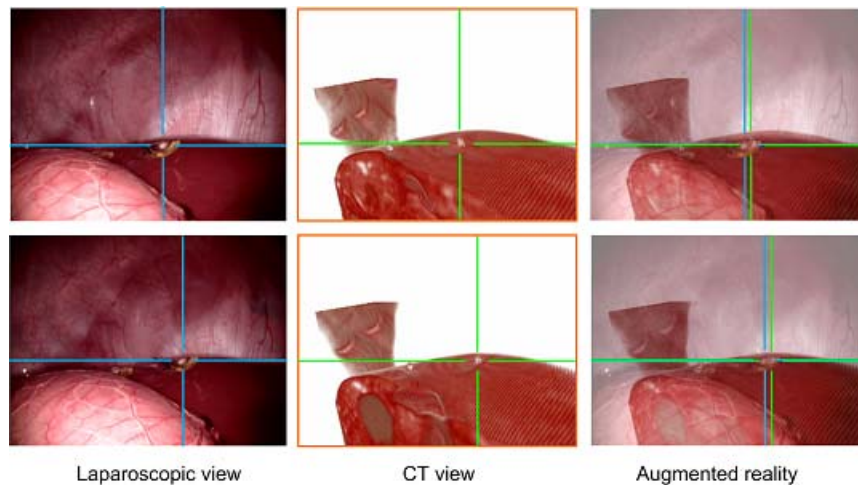
Fig. 7 Live AR leads to a much improved spatial registration between laparoscopic and CT views (right column) from two time instants shown in top and bottom rows. A small residual error can be attributed to experimental errors. The superior accuracy of Live AR is a result of built-in steps for intraoperative motion compensation including breathing.

Discussion

This study tested the feasibility of an ambitious, long-term goal of taking advantage of new developments in volumetric imaging, increasingly being adopted in diagnostic imaging, for enhancing intraoperative visualization during minimally invasive surgeries. Despite significant recent gains in the resolution of the laparoscopes, namely introduction of high-definition [15, 16], their 3D visualization capability remains limited. Essentially a video imaging technique, they cannot reveal structures below the exposed surfaces. The stereo laparoscope—a dual-camera system producing slightly jittered left- and right-eye views for stereopsis—has been another recent attempt to enhance 3D visualization of the surgical field [17]. Although the depth perception is enhanced with these scopes, the fundamental limitation remains. They show only the superficial surfaces and hidden structures and vessels still cannot be uncovered and visualized. AR, as proposed earlier, has provided the missing 3D information but is not accurate for abdominal surgeries, because preexisting CT or MR imaging data employed for 3D visualization may not correctly represent the deformable and changing intraoperative anatomy. The most accurate approach is to perform 3D imaging continuously during the surgery and use the resulting data for AR. We demonstrated here the feasibility of such Live AR approach, whose distinguishing features compared to the conventional approach are summarized in Table 2.

Table 2. Comparison of Live AR and conventional AR

Features	Conventional AR	Live AR
3D Imaging	CT or MR imaging performed once, often preoperatively	Initial CT followed by low-dose CT performed continuously during surgery
Vessel enhancement	From contrast enhancement during preoperative imaging	From substitution of intraoperative CT with contrast-enhanced initial CT
Radiation exposure	Not a concern	Concern addressed by low-dose scanning
Accuracy of AR visualization	Error prone; unable to account for anatomic deformations	Anatomic deformations followed by continuous imaging

Continuous real-time 3D imaging in the OR is the first step to equipping operating surgeons with enhanced visualization capabilities. However, continuous 3D imaging has been technologically difficult until recently. MR imaging remains too slow and significant efforts will be needed to manufacture MR-compatible laparoscopes and surgical instruments. While real-time 3D ultrasonography was recently released [18, 19], its image quality remains suboptimal compared with that of CT and MRI. More important, it cannot image across pneumoperitoneum during laparoscopic surgeries. Multislice CT does not suffer from these problems and can, in fact, image the surgical field several times per second. The 64-slice CT scanner we used could scan a 4-cm section of the body at a high resolution (64 parallel slices of 0.625 mm thickness) approximately once every second. Even newer multislice CT scanners with up to 320 slices can scan 12-cm region with high spatial resolution several times per second [20, 21]. Multislice CT, therefore, was our modality of choice for intraoperative imaging because of its higher speed, higher spatial and temporal resolution, higher volumetric coverage, tool compatibility, and favorable technical development trends.

Two practical challenges with the proposed use of continuous CT are radiation exposure to the patient and surgical team, and the need for administering contrast agents to highlight the

underlying vessels. Deformable image registration that is integral to our imaging protocol addresses both these challenges. The accuracy of image registration is the first and foremost consideration, which was measured to be approximately 1.5 mm. This is acceptable because the primary goal of Live AR is to uncover underlying vessels and features. We also demonstrated the potential of 8-fold reduction in the dose given to the animals in this study by conducting intraoperative CT at 25 mA instead of 200 mA. Further dose savings are possible because the registration accuracy was retained even at 25 mA, but the CT scanner did not allow further lowering the current setting. An earlier study in which a dose simulator software allowed us to create lower-dose scans from a standard-dose CT scan of archived patient images and to explore a larger range of dose settings, including as low as 11 mA, indicated acceptable registration at 11 mA, which represented a 20-fold dose reduction [22]. These observations tell us the potential for additional dose savings through deformable image registration if extrinsic factors do not prevent us from exploring the lower range of dose settings fully.

Visualization of critical underlying structures, especially the vasculature, is important before making surgical dissections. Inherent in our imaging protocol and deformable image registration between initial and intraoperative CT data is a scheme to visualize the vessels without having to use the contrast agent continuously, which is neither permitted nor safe. An advantage of having 3D rendering of the CT data is that one can interact with this view. For example, the surgeon can virtually practice a particular surgical manipulation and observe the effects of it in the CT view before actually making that manipulation. No such interaction is possible with the traditional laparoscopic view. A promising new approach to visualize the vasculature was recently proposed by Crane et al. [23], who exploited tissue oxygenation-based differences in three component images obtained using a three-charged couple (CCD) camera. While this approach needs further testing, a potential drawback is that a lack of true volumetric image of the surgical field unlike in our method will not permit rehearsing a potential surgical manipulation virtually.

The process of Live AR is theoretically the most accurate approach to AR visualization. When compared experimentally, Live AR was indeed found more accurate than conventional AR. Live AR is more accurate because the CT scan used for 3D visualization of the surgical field is acquired exactly when the corresponding laparoscopic image is acquired. The perfect temporal synchronization between the two makes Live AR insensitive to the voluntary and involuntary anatomic changes and resulting CT-laparoscopic overlay free of misregistration. In practice, some minor misregistration was present because of the slow frame rate of CT, finite precision of deformable image registration, and the current manual approach to synchronize CT and laparoscopic imaging systems. Conventional AR, as implemented by us, attempted to superimpose 3D rendering of an intraoperative CT snapshot with laparoscopic images acquired over a period of time. The misregistration in conventional AR was not only higher, but also variable with time. When the breathing phases match, least error is expected. When the two are completely out of phase, maximum error can be expected.

The current study constituted an offline study in that it took many days of data processing before Live AR visualization could be ready. Slow data processing was not a limitation in proving the concept of Live AR, which is a prerequisite to motivate the necessary engineering advances for eventual online implementation of Live AR. A few other limitations were the use of mostly gross and breathing-induced anatomical changes to simulate intraoperative changes. More realistic

surgical moves could not be tested because they required an operator to perform those while standing next to a CT scanner with the scanner on, when the feasibility of continuous low-dose CT was still being investigated. This limitation can be overcome in the future when it is possible to reduce the dose even further, as suggested by our results here. A lack of integration between the CT scanner and the PC controlling the optical tracker and laparoscopic video frame grabber was also a limitation that necessitated a manual temporal synchronization between the intraoperative CT scans and laparoscopic video. The use of a 0-degree laparoscope and not allowing camera rotation were limitations that did not interfere with the feasibility testing of Live AR and can be overcome rather easily in the future. Only 4-cm coverage and approximately 1 Hz refresh rate were also limitations that will get addressed by newer CT scanners with more slices and faster rotation time.

We believe we succeeded in proving the feasibility of Live AR, so it is imperative to consider future efforts needed for making continuous CT-guided laparoscopic surgery and Live AR routine. First, the CT technology needs to be improved in many ways. The scanner we used as well as most current scanners cannot reconstruct 64 slices per second needed during continuous imaging. The required reconstruction speed will grow higher for newer scanners with more than 64 slices and faster rotation capability. The scanners also need to provide real-time access to the reconstructed images, a capability that does not exist currently. Yet another enhancement would be the ability to lower the x-ray dose to extremely low levels that are currently not permitted. Second, image registration needs to be made even faster. The typical time of image registration for 64-slice data currently is approximately 1.5 min, whereas one new registration per second needed to be performed in our current experiments. Third, a much tighter integration among many systems and subsystems is needed. These include the CT scanner, the laparoscope and surgical tools, the optical tracker, image registration module, 3D visualization workstation, etc. Some other technical improvements for the final implementation will include a redesign of the surgical tools to minimize metal artifacts in CT and more robust calibration devices and procedures.

In conclusion, our work combines emerging continuous 3D CT imaging with minimally invasive laparoscopic surgery for improved intraoperative visualization which we have called Live AR. Continuous low-dose CT of the dynamic surgical field at safe and acceptable radiation doses and using high-speed deformable image registration to generate diagnostic-quality contrast-enhanced CT images of the intraoperative anatomy will enable high-quality 3D visualization of the surgical field during laparoscopy. We have successfully demonstrated the initial feasibility of this concept, which, with further technical enhancements, could be made routine. Live AR promises to lead to improved precision in laparoscopic surgeries with fewer complications. Aided by improved visualization, it is also expected that many surgeries that are currently performed in an open invasive fashion can instead be performed minimally invasively, expanding the benefits of minimally invasive surgeries to more patients and during more procedures.

Acknowledgments

1. Himal HS (2002) Minimally invasive (laparoscopic) surgery. Surgical Endoscopy 16:1647-1652
2. Rosen M, Ponsky J (2001) Minimally invasive surgery. Endoscopy 33:358-366
3. Osborne DA, Alexander G, Boe B, Zervos EE (2006) Laparoscopic cholecystectomy: past, present, and future. Surg Technol Int 15:81-85
4. Hanly EJ, Talamini MA (2004) Robotic abdominal surgery. Am J Surg 188:19S-26S
5. Harrell AG, Heniford BT (2005) Minimally invasive abdominal surgery: lux et veritas past, present, and future. Am J Surg 190:239-243
6. Fuchs H, Livingston MA, Raskar R, Colucci D, Keller K, State A, Crawford JR, Rademacher P, Drake SH, Meyer AA (1998) Augmented Reality Visualization for Laparoscopic Surgery. In Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes In Computer Science 1496:934-943
7. Marescaux J, Rubino F, Arenas M, Mutter D, Soler L (2004) Augmented-reality-assisted laparoscopic adrenalectomy. JAMA 292:2214-2215
8. Mutter D, Bouras G, Marescaux J (2005) Digital technologies and quality improvement in cancer surgery. Eur J Surg Oncol 31:689-694
9. Walimbe V, Shekhar R (2006) Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations. Med Image Anal 10:899-914
10. Dandekar O, Shekhar R (2007) FPGA-accelerated deformable image registration for improved target-delineation during CT-guided interventions. IEEE Trans Biomed Circuits Syst 1:116-127
11. Dandekar O, Castro-Pareja C, Shekhar R (2007) FPGA-based real-time 3D image preprocessing for image-guided medical interventions. Journal of Real-Time Image Processing 1:285-301
12. Feuerstein M, Mussack T, Heining SM, Navab N (2008) Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection. IEEE Trans Med Imaging 27:355-369
13. Shahidi R, Bax MR, Maurer CR, Jr., Johnson JA, Wilkinson EP, Wang B, West JB, Citardi MJ, Manwaring KH, Khadem R (2002) Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. IEEE Trans Med Imaging 21:1524-1535
14. Bouguet J-Y (2009) Camera Calibration Toolbox for Matlab. Available at: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html. Accessed 28 May 2009.
15. Hagiike M, Phillips EH, Berci G (2007) Performance differences in laparoscopic surgical skills between true high-definition and three-chip CCD video systems. Surg Endosc 21:1849-1854
16. Pierre SA, Ferrandino MN, Simmons WN, Fernandez C, Zhong P, Albala DM, Preminger GM (2009) High definition laparoscopy: objective assessment of performance characteristics and comparison with standard laparoscopy. J Endourol 23:523-528
17. Miller A, Allen P, Fowler D (2004) In-vivo stereoscopic imaging system with 5 degrees-of-freedom for minimal access surgery. Stud Health Technol Inform 98:234-240
18. Cannon JW, Stoll JA, Salgo IS, Knowles HB, Howe RD, Dupont PE, Marx GR, del Nido PJ (2003) Real-time three-dimensional ultrasound for guiding surgical tasks. Comput Aided Surg 8:82-90

19. Sugeng L, Weinert L, Thiele K, Lang RM (2003) Real-time three-dimensional echocardiography using a novel matrix array transducer. Echocardiography 20:623-635
20. Kalender WA (2006) X-ray computed tomography. Phys Med Biol 51:R29-43
21. Rogalla P, Kloeters C, Hein PA (2009) CT technology overview: 64-slice and beyond. Radiol Clin North Am 47:1-11
22. Dandekar O, Siddiqui J, Walimbe V, Shekhar R (2006) Image registration accuracy with low-dose CT: how low can we go? In Proceedings of 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 502-505
23. Crane NJ, McHone B, Hawksworth J, Pearl JP, Denobile J, Tadaki D, Pinto PA, Levin IW, Elster EA (2008) Enhanced surgical imaging: laparoscopic vessel identification and assessment of tissue oxygenation. J Am Coll Surg 206:1159-1166

An image registration–based approach for continuous volumetric CT-guided interventions

Raj Shekhar, Ananthranga Prithviraj

Department of Diagnostic Radiology, University of Maryland School of Medicine

Purpose

Minimally invasive image-guided interventions (IGIs) that include biopsies, ablations, and surgeries are less than optimal because of the unavailability of continuous three-dimensional (3D) visualization of the anatomy. When continuous or real-time, intraoperative imaging remains two-dimensional as in conventional and computed tomography (CT) fluoroscopy, ultrasound, magnetic resonance (MR) imaging, and endoscopy. When three-dimensional (3D), as in volumetric CT and MR imaging, the imaging remains temporally discrete and the resulting image guidance is stop-and-go and inefficient.

Recent advances in multidetector CT (MDCT) are beginning to permit continuous 3D imaging during an IGI. With the latest MDCT scanners, it is now possible to scan up to 10–12-cm thick regions of the anatomy at an extremely high spatial resolution multiple times per second. But radiation exposure concerns and inability to visualize the vasculature (contrast agents cannot be administered continuously) limit clinical implementation of continuous 3D CT, despite being technically feasible.

We present here a novel concept based on high-speed 3D image registration that addresses both these problems. We acquire a single contrast-enhanced volumetric CT scan (called initial CT) at a diagnostic dose at the start of the IGI. Subsequently, CT is operated at a low dose without contrast. A diagnostic-quality contrast-enhanced image of the operative field is obtained by rapidly and nonrigidly registering the initial CT with intraoperative low-dose CT. We present here the feasibility of our concept in terms of registration time and accuracy, savings in radiation dose, and intraoperative vessel visualization.

Methods

Our imaging specimen was a swine prepared for a mock laparoscopic liver surgery under experimental CT guidance. All CT images were acquired using a 64-slice CT scanner (Philips Brilliance-64) following pneumoperitoneum. Before imaging, 4 markers (2-4 mm guidewire pieces) were implanted in the liver parenchyma and 2 2.3-mm calcium markers sutured onto the liver surface for objective validation of image registration. The initial CT was a helical CT scan of the liver (53-cm axial coverage) at normal breathing with arterial phase enhancement at a diagnostic dose (250 mAs tube current). The swine was then overventilated to accentuate liver motion and deformation from the time of initial CT. CT scanning, simulating intraoperative imaging, was then performed at high, medium, and low doses (200, 75, and 25 mAs, respectively) to determine the lower limit of low-dose CT. For all 3 doses, this CT was performed in 2 modes: helical and axial. The helical mode allowed complete coverage of the liver but provided only a snapshot of it. The axial mode could acquire repeated scans (we acquired 100) at 0.9 Hz, but the 4-cm axial coverage of the CT scanner used permitted only partial liver coverage. Using a previously reported, hardware-accelerated implementation of nonrigid image registration, we registered the initial CT with each of the 3 helical CT scans and each of the 100 scans in the 3 axial CT scan sequences. The initial relative position of image pairs was based on slice location data saved with images. Using the implanted markers, the initial and postregistration misalignments (reflective of registration accuracy) were computed. The time of registration was also recorded.

Results

For 3 helical scans, approximately 3-mm initial misalignment reduced to approximately 1.5 mm for all 3 doses after nonrigid registration of initial CT with intraoperative CT (Table 1). The results show that the dose had virtually no effect on registration accuracy, indicating that intraoperative CT could be performed at 25 mAs. Most structural mismatches before registration were removed after registration (Figure 1). In Figure 2, volume rendering of the original intraoperative CT and registered initial CT (representing intraoperative anatomy) is shown. Note that using our concept, the vasculature can be visualized throughout an IGI without having to administer CT contrast. A similar registration was performed between initial CT and axial scan sequences at the 3 doses. The mean initial and postregistration misalignments (averaged over 100 scans) are shown in Table 2. The nonrigid registration exhibited acceptable accuracy despite small coverage. These results, too, suggest that intraoperative CT at 25 mAs is acceptable. The mean registration time for larger helical data sets was 430 s and for smaller axial scans was 63 s.

Table 1. Image misalignment before and after registration for helical scans.

Dose (mAs)	Initial Misalignment (mm)	Misalignment after Registration (mm)
200 (high)	3.12	1.47
75 (medium)	3.63	1.67
25 (low)	3.25	1.45

Table 2. Average image misalignment before and after registration for axial scans.

Dose (mAs)	Mean Initial Misalignment (mm)	Mean Misalignment after Registration (mm)
200	4.40	1.05
75	4.51	1.12
25	4.69	1.21
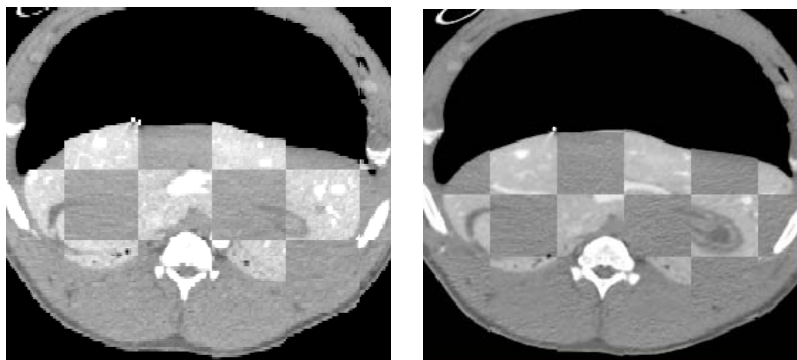
Figure 1. Superposition of initial CT and intraoperative CT (acquired at 200 mAs) before (left) and after (right) nonrigid registration. Note better structure alignment after registration.
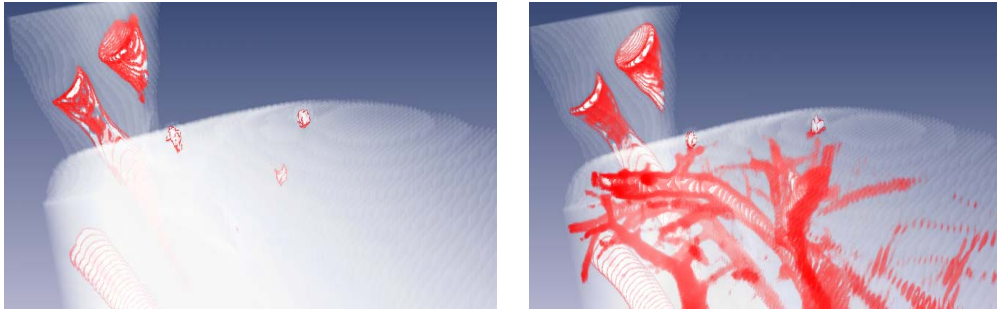
Figure 2. Volume rendering of original intraoperative CT and registered initial CT. Note that the latter shows the vasculature.

Conclusions

We have presented proof-of-concept results using high-speed image registration to substitute noncontrast low-dose intraoperative CT images with a modified contrast-enhanced diagnostic-quality initial CT image during an IGI. The latter contains contrast enhanced structures for which visualization may be critical during an IGI. Our strategy also led to a 10-fold savings in radiation dose (tube current reduced from 250 mAs to 25 mAs). 25 mAs was the lowest setting on the CT scanner used, suggesting that further dose savings are possible if the CT scanner could be operated at a lower dose. Finally, the residual misregistration on the order of 1 mm is acceptable, because the targeting uncertainty in most IGIs is currently much greater. It should also be noted that the concept extends to *any* preoperative image and is likely to vary depending on specific imaging needs of an IGI. Ours was an offline feasibility study. Its clinical implementation will require further speed improvement of CT reconstruction and image registration. Overall, we have presented a novel concept along with demonstration of its feasibility that promises to enable use of the latest MDCT for continuous 3D visualization at acceptably low doses during most IGIs.

Mapping the Way to a Dual Display Framework for Laparoscopic Surgery

Ruigang Yang, C. Melody Carswell, Xianwang Wang, Qing Zhang, Qiong Han, W. Brent Seales, and Cindy H. Lio

University of Kentucky, Departments Of Computer Science and Psychology, and Center for Visualization and Virtual Environments.

Abstract – Many performance and workload problems associated with the use of traditional laparoscopic displays are the result of spatial disorientation. This premise has guided our development of a dual display framework for computer-augmented surgical displays, allowing us to take guidance from research on how to design successful navigation aids (navaids) for large-scale environments. Our dual-display combines the traditional scope (forward track) view with a computationally-generated global 3D (map) view. The latter provides a wider field of view, explicit cues to depth and scale, and a way to view interior and exterior surfaces of target anatomy from different approach angles. One way to implement such a 3D view is to extract images of surface textures from a laparoscopy video sequence and then map the texture onto pre-built 3D objects, for example surface models derived from MR/CT. We describe an algorithm that takes advantage of the fact that nearby frames within a video sequence usually contain enough coherence to allow 2D-2D registration, a much better understood problem than 2D-3D registration. Our texturing process can be bootstrapped by an initial 2D-3D manual-assisted registration of the first video frame followed by mostly-automatic texturing of subsequence frames. Initial research on the validity of our technical approach indicates that it improves registration performance compared to a standard registration technique that relies on camera tracking. Ongoing technical and usability evaluations of the system are being conducted in order to ensure system functionality.

Introduction

When surgeons view the surgical field through a typical laparoscope, they are faced with visual images that are degraded in a variety of ways. Compared to the view of target anatomy afforded in open procedures, the field of view is reduced and stereoscopic depth cues are eliminated. To further complicate matters, relative locations and movement trajectories may be misperceived because the surgeon's spatial frame of reference may not be aligned with those of the laparoscope and the display screen. Although haptic information can sometimes help disambiguate degraded visual cues such as these, haptic information is also reduced during most minimally invasive procedures.

Researchers have found that the reduction of perceptual cues such as those described above have reliable negative effects on simulated surgical performance (e.g., DeLucia, Mather, Griswold, & Mitra, 2006: Emam, Hanna, & Cuschieri, 2002) as well as on mental workload and stress (e.g., Klein, Warm, Riley, Matthews, Gaitonde, and Donavan, 2008; Klein, Riley, Warm, &

Matthews, 2005). Although surgeons can sometimes learn to adapt to these challenges, it may be at the cost of increased training and greater mental effort. Thus. the better design of surgical displays is a pressing need as minimally invasive procedures become more prevalent. Computer-augmented surgical displays – so-called "smart" images – are a natural approach because they can provide the flexibility to manipulate digitized images, replacing lost sensory cues, correcting distorted ones, and even providing additional cues not available in open surgery. Such augmentation has been seen as key to the next generation of surgical displays (e.g., Pulli et al. 1997, Fuchs el al. 1998, Paul et al. 2005). In this paper, we describe our approach to the development of one such system.

Surgical Displays as Navaids. In specifying the requirements of computer-augmented surgical displays, we have found it useful to conceptualize many of the problems encountered during laparoscopic surgeries as variations on a single theme. That is, many of the negative consequences of impoverished imagery really fall into one general class of cognitive failures -- *spatial disorientations*. Many surgical tasks can be thought of as navigation or wayfinding tasks including, for example, route planning, instrument steering, landmark identification, and obstacle avoidance. This conceptualization is consistent with the research and views of Cao and Milgram (2000) and Summers (1997), who brought attention to the ease with which surgeons can become lost within the confines of the human body as readily as pedestrians can become lost in an unfamiliar city. And, just as navigational aids (navaids) can be designed to help the pedestrian reach his or her goal more efficiently and safely, surgical displays can be augmented with task-appropriate information that can also show similar benefits.

We approached the problem of helping surgeons avoid "getting lost" by first looking at cognitive science research pertinent to the development of navaids for use in large-scale environments. In their extensive review of this literature, Taylor, Brunye, and Taylor (2008) describe a variety of barriers to efficient and accurate navigation. Among these are 1) a restricted field of view, 2) occlusions, 3) incongruent frames of reference between navaids and the terrain being traversed, 4) ambiguities of scale, and 5) lack of redundancy across perceptual modalities. Clearly, many of these general barriers reflect the reality of most minimally invasive surgical landscapes. The list therefore serves as the inspiration for our *dual display framework* shown in Figure 1.

In the dual display system, the traditional scope (forward track) view shown on the right side of Figure 1 is augmented with a computationally-generated global 3D (map) view on the left. The global view provides a wider field of view with explicit depth information for both the exterior and interior of target anatomical objects. The global viewpoint is scope-independent and can be manipulated in a variety of ways, for example allowing the surgeon to "see through" structures on demand. The size of objects can be made unambiguous by the application of a fixed-size grid to structures of interest, again on demand. In addition, the relationship between the orientation of the detailed scope view and the current global view is made explicit by either 1) highlighting the area of the global view that is the focus of the scope view (as in Figure 1)

when the global and scope view are otherwise aligned, or 2) showing a scope icon on the global view to indicate the spatial relationship between the two views when they are not aligned.
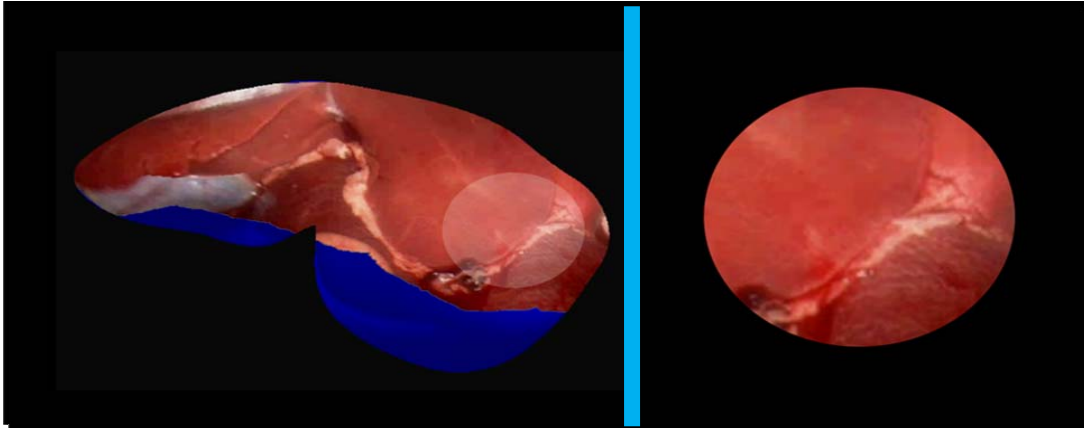

Figure 1. The dual-display framework. The right image shows the captured endoscope view. The left image shows an enhanced 3D view in which the scope view is registered with its corresponding 3D model.

Technical Challenges. Although the dual display would, in principle, address many of the barriers to efficient surgical wayfinding described above, its implementation involves a number of technical challenges.   One of the biggest obstacles involves registration – the accurate spatial integration of information from multiple sources. Registration between data acquired under different modalities such as video and CT has long been an open problem in medical imaging. Although there are many algorithms for 2D to 2D registration, 2D to 3D registration is far more challenging.   The approach we take is to combine a 3D anatomical object and its corresponding 2D laparoscopic view by treating the camera video sequence as texture information for the 3D object.

A Method for Mapping Texture to Geometry

Traditionally, geometry and texture are acquired at the same time with the same sensor [Sako and Fujimura 2000; Dey et al 2002]. In this case, the images are already aligned to the model and no further 3D-2D registration is needed.   However, in most cases, specialized 3D scanners are used to acquire the precise geometry, and high quality digital cameras are used to capture detailed texture information. Thus, the images have to be registered with the 3D geometry to build correspondences between the geometry and texture information.

This texture-to-geometry registration problem can be handled by tracking the camera. With precise instrumentation and/or fiducial markers, satisfactory results have been obtained. However, there are certain applications in which neither of these requirements can be satisfied. For example, in a surgical setting, the laparoscope camera cannot be directly tracked. The tracking sensor can only be attached to the outside of the scope, and the long offset between the sensor and the tracker magnifies the tracking error. Given an endoscope's narrow field of

view, a prohibitively high density of fiducial markings would also be needed.   Another approach for texture mapping is to find parameterizations of a 3D geometric model onto a 2D texture domain, while maintaining certain criteria, such as the minimization of distortions and compliance with user-defined feature point locations (e.g., [Levy 2001, Desbrun et al. 2002, Kraevoy et al. 2003]. While this works reasonably well for a single image, applying the manual selection of feature points in every video frame is too time-consuming and tedious for implementation in the operating room.

To realize our dual-display framework, we combine visual tracking with parameterization-based texture mapping. The basic idea is to boot-strap the registration process with a user-interactive method (such as the one from Levy, 2001) and then use vision-based tracking to find 2D-2D sparse correspondences between images. Since the first image has 2D-3D correspondences, each new image, based on the overlapping area to the previous image, can be "pasted" onto the 3D model automatically using a parameterization based method.   In this way, we convert the difficult 3D-to-2D registration problem into the well studied and better understood problem of 2D-to-2D matching. The method can be used for images that cannot be modeled by perspective projection. In addition, when the camera is indeed projective and the object is (almost) rigid, we introduce a projective correction term in the parameterization process so that accurate registration can be achieved over a wide range of viewpoint changes. Unlike projective texture mapping, this projective correction term is a soft constraint so our method is more robust against errors in the 3D model or even slightly deformed models. In both cases, we avoid the problem of camera calibration or external tracking.

In essence, our approach combines the strength of both visual tracking (automatic) and parameterization-based texture mapping (more flexible), leading to a new means to quickly and semi-automatically add textures to 3D models.   This provides the technical foundation for our dual display system.

Summary of Technical Approach. The key to our dual-display framework is the registration of 2D video endoscopic images onto the 3D object's surfaces. Our algorithm takes advantage of the fact that nearby frames within a video sequence usually contain enough coherence to allow a 2D-2D registration – the "stitching" of one video frame to another to form a panorama . Thus, the texturing process can be boot-strapped by an initial 2D-3D manually-assisted registration of the first video frame followed by mostly-automatic texturing of subsequent frames. Figure 2 shows our pipeline for the texturing process. It includes three stages: single view mapping, panorama construction, and incremental texture mapping.

1)   *Single view mapping*: we adopt the Least Square Conformal Mapping (LSCM) algorithm [Levy 2001], where a user can assign correspondences between a 3D model and a 2D texture map and the system optimizes a mapping that minimizes distortions. It consists of three components: feature correspondences, parameterization, and linear system solver.

2)   *Panorama construction*: a series of endoscopic images are taken from different viewpoints and stitched into a single large image that is continuous in geometry and shading.

3)   *Incremental texture mapping*: this is the heart of our algorithm. The endoscopic images except for the first frame from the video sequence are mapped onto the geometry incrementally. The system propagates the user constraints defined in the single view mapping to the panorama. In this way, there is less user interaction required. The user can add new correspondences whenever required.
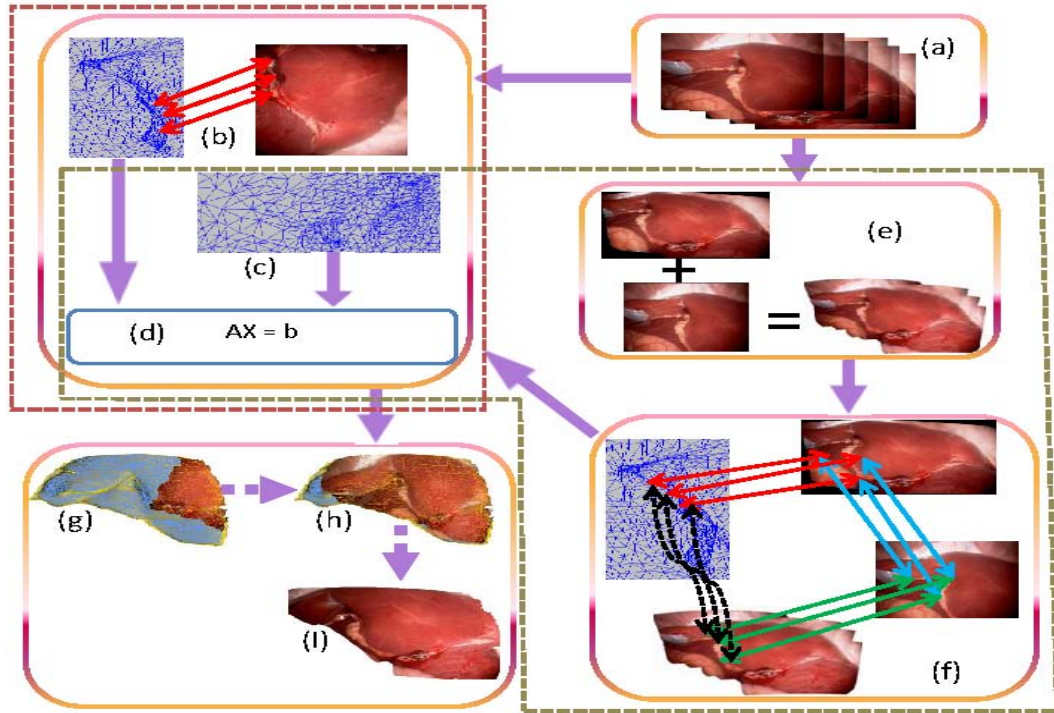
Figure 2 The pipeline of our system for texturing endoscopic images to 3D surfaces. (a)images extracted from the endoscopic video. (b)3D-2D feature correspondences; (c)parameterization; (d)linear system solver; (e)panorama construction; (f)incremental texturing mapping; (g)the result of the single view mapping; (h) the immediate result of texture mapping; (h) The final result of texture mapping.

    Sample of Texture Mapping Results. Based on the approach presented above, we have implemented a flexible system for texture mapping. It is an interactive system, where the user can add new frames from video sequences and edit the correspondences between the model and texture images. Three examples of varying complexity are presented to demonstrate the use of our system. Figure 3, Figure 4, and Figure 5 show the 3D model, the intermediate registration result, and a view of the final texture mapping.
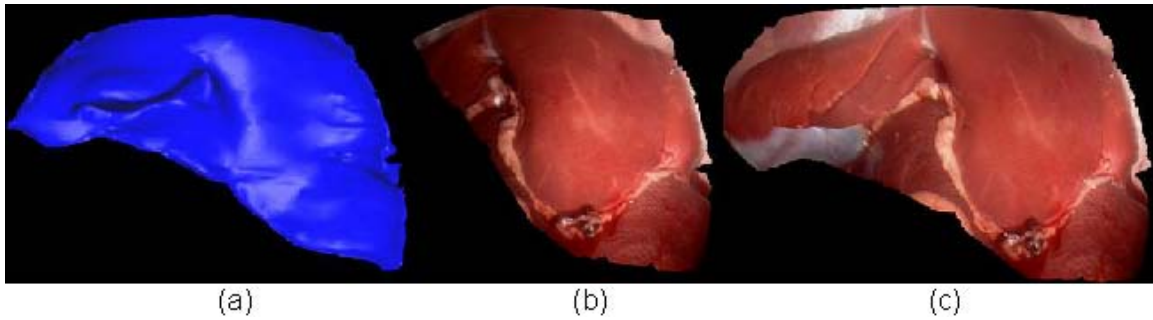
Figure 3. Registration result of real data, pig liver. (a) liver model; (b) intermediate registration result; (c) final registration result
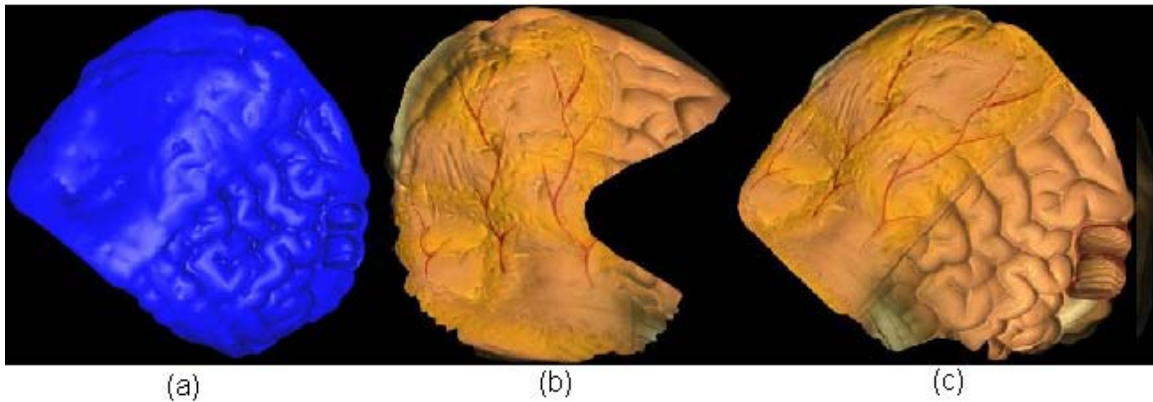

Figure 4. Registration result of phantom data, human intestine. (a) the front of human intestine model; (b) intermediate registration result; (c) final registration result
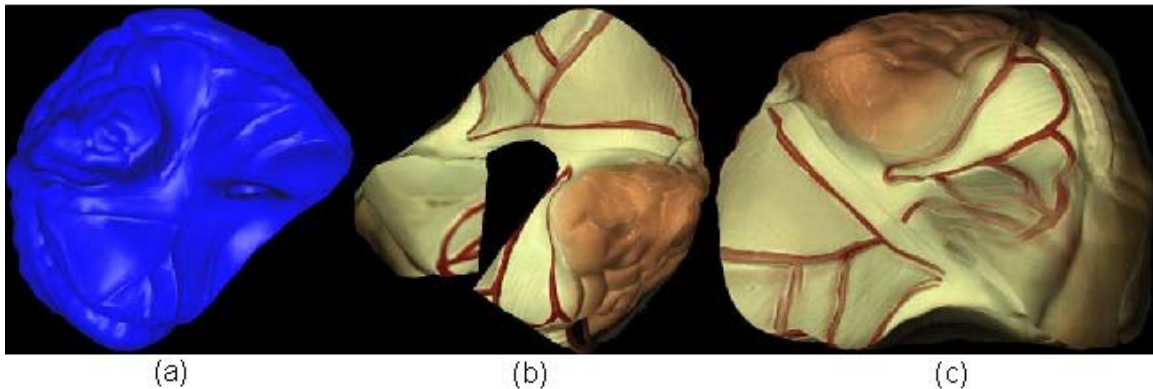

Figure 5. Registration result of phantom data, human intestine. (a) the back of human intestine model; (b) intermediate registration result; (c) final registration result

Visualization

One potential problem users will encounter with our dual-display framework
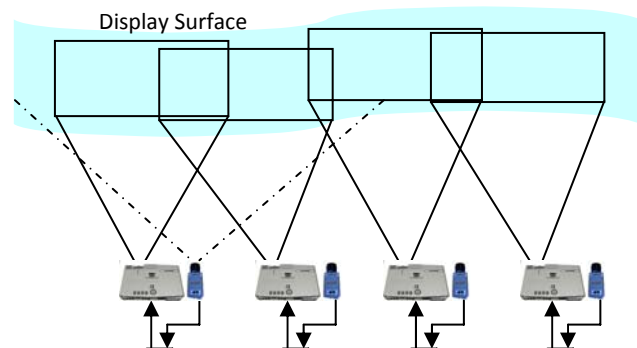
Figure 6. A schematic of the display system architecture. Each projector is augmented with a camera with a wider field view. Each pair is connected to a computing platform (e.g. a PC) this is networked.

is that it will require more screen real estate than conventional surgical displays. This is because it will show both the scope view and the 3D view in high resolution. Toward this end, we leverage our previous work on multi-projector displays [Yang et al. 2001]. That is, we will use a cluster of projectors to create a large seamless, high-resolution display (shown in Figure 6 and 7). Unlike traditional display clusters that rely on manual mechanical alignment, we have developed camera-based calibration techniques that can align a casually placed projector array in a matter of minutes [Brown et al. 2006]. That significantly reduces the requirement for space and maintenance and makes it possible to use a projector array in an already crowded OR.

Figure 7. A prototype six-projector display in our visualization lab. The overlap among the images is made visible in order to show the contribution of the individual projectors.

Future Work

In this paper we have presented a vision of a dual-display framework in which the surgeon can see not only acquired endoscopic imagery but also computationally-enhanced views with proper 3D cues. We are facing the technical challenges required to implement such a display system because we believe that by providing a global view of the surgical site we will enhance the surgeon's ability to navigate within the body and will reduce spatial disorientation. However, this outcome is not a foregone conclusion. It is possible that having more than one view of the surgical field could result in too much competition for the surgeon's attention. And it is possible that users might only attend to one view and ignore the other. DeLucia, Hoskins, and Griswold found evidence that this was the case when they explored the potential benefits of providing concurrent views from three scopes in order to increase the ability of users to judge depth when performing a simple laparoscopic training tasks. These authors found no advantage of having multiple viewpoints over a single view, but they suggest that this is because their research participants tended to focus on only one of the available visual channels. If the three

orientations were integrated into a single global view, as proposed here, then perhaps a performance advantage would be achieved.   Questions of this sort, as well as questions regarding the best way to allow surgeons to manipulate the global (map) view in the dual display system, require careful user testing as part of total system evaluation.

References

[Brown et al 2005] Michael S. Brown, Aditi Majumder, Ruigang Yang, Camera-Based Calibration Techniques for Seamless Multi-Projector Displays, in IEEE Transactions on VISUALIZATION AND COMPUTER GRAPHICS (TVCG), 11(2): 193-206, 2005

Cao, C. G. L., & Milgram, P. (2000). Disorientation in minimal access surgery: A case study. In Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society (pp. 4.169–4.172). Santa Monica, CA: Human Factors and Ergonomics Society.

[Clarkson et al. 1999] D. L. H. Matthew J. Clarkson, Daniel Rueckert and D. J. Hawkes, "Registration of multiple video images to preoperative ct for image-guided surgery," in SPIE Medical Imaging, 1999, pp. 14–23.

DeLucia, P.R., Hoskins, M.L., & Griswold, J.A. (2004). Laparoscopic Surgery: Are Multiple Viewing Perspectives better than One?   In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (pp. 1661-1664), Santa Monica, CA: Human Factors and Ergonomics Socieity.

DeLucia, P. R., Mather, R. D., Griswold, J. A., & Mitra, S. (2006). Toward the improvement of image-guided interventions for minimally invasive surgery: Three factors that affect performance. Human Factors, 48, 23–38.

[Desbrun et al. 2002] M. Desbrun, M. Meyer, and P. Alliez, "Intrinsic parameterizations of surface meshes," in Proceedings of Eurographics, 2002.

[Dey et al 2002] D. Dey, D. G. Gobbi, P. J. Slomka, K. J. M. Surry, and T. M. Peters, "Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: Creating stereoscopic panoramas," pp. 23–30, 2002.

Emam, T. A., Hanna, G., & Cuschieri, A. (2002a). Comparison of orthodox versus off-optical axis endoscopic manipulations. Surgical Endoscopy, 16, 401–405.

[Fuchs et al 2003] H. Fuchs, M. A. Livingston, R. Raskar, D. Colucci, K. Keller, A. State1, J. R. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer, "Augmented reality visualization for laparoscopic surgery," in MICCAI'98, LNCS 1496, 1998, pp. 934–943.

Klein, M. I., Riley, M. A., Warm, J. S., & Matthews, G. (2005). Perceived mental workload in an endoscopic surgery simulator. In Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting (pp. 1014–1017). Santa Monica, CA: Human Factors and Ergonomics Society.

Klein, Martina I.; Warm, Joel S.; Riley, Michale A.; Matthews, Gerald; Gaitonde, Krishnanath; Donovan, James F. (2008). Perceptual Distortions Produce Multidimensional Stress Profiles in Novice Users of an Endoscopic Surgery Simulator. Human Factors, 50(2) 291-300.

[Kraevoy et al. 2003] V. Kraevoy, A. Sheffer, and C. Gotsman, "Matchmaker: Constructing constrained texture maps," in Proceedings of SIGGRAPH 2003, 2003, pp. 326–333.

[Levy 2001] B. Levy, "Constrained texture mapping for polygonal meshes," in Proceedings of SIGGRAPH 2001, 2001, pp. 417–424.

[Paul et al. 2005] P. Paul, O. Fleig, and P. Jannin, "Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery: Methods and performance evaluation," vol. 24, no. 11, pp. 1500–1511, 2005.

[Pulli, et al. 1997] K. Pulli, M. Cohen, T. Duchamp, H. Hoppe, L. Shapiro, and W. Stuetzle, "View-based rendering: Visualizing real objects from scanned range and color data," in Eurographics Rendering Workshop. Springer Wien, June 97, pp. 23–34.

[Sako and Fujimura 2000] Y. Sako and K. Fujimura, "Shape similarity by homotropic deformation," The Visual Computer, vol. 16, no. 1, pp. 47–61, 2000.

Summers, R. M. (1997). Computers in radiology: Navigational aids for real-time virtual bronchoscopy. American Journal of Roentgenology, 168, 1165–1170.

Taylor, H.A., Brunye', T.T., & Taylor, S.T. (2008). Spatial Mental Representation: Implications for Navigational System Design. In C.M. Carswell (Ed.), Reviews of Human Factors and Ergonomics, Vol. 4. Santa Monica, CA: Human Factors and Ergonomics Socieity.

[Yang et al. 2001] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. S. Brown. PixelFlex: A Reconfigurable Multi-Projector Display System. In *Proceeding of IEEE Visualization 2001*, pages 167–174, San Diego, CA, 2001

Using Formal Qualitative Methods to Guide Early Development of an Augmented Reality Display System for Surgery

C. H. Lio[1], C. M. Carswell[1, 2], Q. Han[1], A. Park[5], S. Strup[3], W. B. Seales[1], D. Clarke[1],
G. Lee[5], and J. Hoskins[4]

[1] Center for Visualization and Virtual Environments, [2] Department of Psychology,
[3] Department of Urology, and [4] Minimally Invasive Surgical Center, University of Kentucky
[5] Department of Surgery, University of Maryland

Nine laparoscopic surgical experts (2 residents, 4 fellows, and 3 surgeons) underwent semi-structured interview questions to evaluate the concept of a "dual-view" display for laparoscopic surgery. The 30-40 minute audio-recorded interviews were transcribed, submitted to an open source qualitative program for classification and categorizing, and were condensed for the iterative processes of analysis and interpretation. Findings revealed that despite the relatively brief interview sessions and limited number of surgical experts available, the experts provided sufficient insights and suggestions to guide further development of prototypes. This means that the use of semi-structured interviews as an expert knowledge elicitation technique may be suitable for assessing the development of augmented reality display systems for surgical and training applications, and it may have promise for the development of augmented and virtual environments more genially.

INTRODUCTION

During laparoscopic surgeries, a surgeon's view of the patient's anatomy is limited to the monitor's images projected from a laparoscopic camera that is inserted into a small incision on the patient. This conventional 2D laparoscopic surgical visualization system impairs surgeons' depth perception and eye-hand coordination. Not surprisingly, there is growing interest in better supporting this challenging surgical approach, and this has led to various technological developments and integrative surgical innovations.

This technological trend begs the question of whether these "apparent advances are, in fact, real" (Carswell, Clarke, & Seales, 2005, p.80). That is, do novel visualization systems, including augmented reality (AR) technologies, stereoscopic displays, and both photorealistic and nonphotorealistic rendering really support laparoscopic surgical performance? Are these systems truly user-centered designs? Have laparoscopic surgeons and trainees been an integral part of these technologies' developmental process? And what methods can best be used to obtain feedback from surgical experts during the conceptual design of such systems?

Most laparoscopic technology evaluations have focused on validating fully developed interventions that have already been adopted or are about to be integrated in training programs and in operating rooms (e.g. Felsher, et al., 2005; Nguan, Girvan, & Luke, 2008, Stefanidis, et al., 2007). Limited studies have investigated the role of laparoscopic experts in the early development of new visualization technologies, although there have been discussions about the importance of their inputs (e.g., Swanstrom, Whiteford, & Khajanchee, 2008). For seamless deployment of novel visual augmentations to laparoscopic surgical practice, it appears that systematic methods of involving laparoscopic experts during the developmental stages are necessary. This goal, however,

may be difficult to achieve when the target users of a system are largely unavailable for all but the briefest interactions with designers. Most typical knowledge elicitation methods for complex domains such as surgery are usually time intensive with multiple sessions (e.g. focused and structured observation participation).

The present study examined the suitability of abbreviated expert knowledge elicitation methods for involving laparoscopic surgeons in the early development of a visualization prototype. Here, "expert" is loosely referred to as laparoscopic surgeons, fellows, and residents. Expert knowledge elicitation, under the larger process of knowledge acquisition (KA), consists of a range of techniques that collect a domain expert's knowledge and problem-solving cognitive processes (Cooke, 1994; Shadbolt and Burton, 1995). The goal of these techniques is to transform the acquired knowledge into a model that emulates an expert's skill-, rule-, and knowledge-based (S-R-K) behaviors (Rasmussen, 1986).

There are three main families of knowledge elicitation techniques: 1) observations and interviews, 2) process tracing methods, and 3) conceptual techniques (Cooke, 1994). Each of these families is further divided into classifications of procedures to meet the array of situations and human-computer interface development goals encountered in practice.

Knowledge elicitation techniques have traditionally been practiced in cognitive research but have become widely accepted in fields such as education, anthropology, training, marketing, and knowledge management (KM) (Jetter, Schroder, Kraaijenbrink, and Wijbhoven, 2006). Their application has also been well received in knowledge engineering for the development of knowledge-based systems (KBS) (e.g., MESICAR, a rheumatology diagnostic support system; Horn, 1989). These systems' development often consists of identification of system components and relationships (e.g., Vennix and Gubbels, 1992), problem

description, model structure conceptualization, and model boundary parameterization (Vennix, Anderson, Richardson, and Rohrbaugh, 1992).

Similar information will almost certainly be pertinent to the development of surgical technologies, especially in terms of placing the design and engineering of the innovation within context. For instance, surgeons may provide ideas for new information to display as well as new visualization characteristics that might aid the use or comprehension of the displayed information. They may also describe problematic features to avoid and identify constraints of certain surgical procedures in relation to the innovation.

In this study, we wanted to find whether expert knowledge elicitation is suitable for guiding our development of an augmented surgical display system to support minimally invasive surgery. Specifically, we used the elicitation techniques of semi-structured interviews and prototyping to gain insights from laparoscopic surgeons and trainees about our "dual-view" display, so named because it provides both the real, local surgical window of video images from the laparoscope and a "global view" of target anatomy within the context of neighboring structures and functional systems.

Semi-structured interview offered the flexibility to raise additional questions during the interview while having prepared questions to guide the general course of the session. Our immediate goal was to collect feedback from the surgical experts to advance our development of the dual-view display. The second goal was to explore the feasibility of expert knowledge elicitation as a systematic method in assessing the early development of surgical display innovations.

"Dual-View" Display

The initial development of the dual-view display concept was guided by cognitive ergonomics principles such as 1) exploitation of redundancy, context, and expectancy, 2) reduction of information access effort, and 3) reduction of memory loads. It involves a visualization integration technique based on the goal of computational efficiency, which is necessitated by the ultimate desire to provide these images in real time with minimum lag in response to surgeons' inputs. The dual-view display registers the original camera view onto pre-built 3D (m-rep) shape models in one of three ways, each method differing in the level of visual integration provided to the user. In the "integrated" dual-view display, the camera view is embedded in its approximate location on a panorama created by "stitching" or "mosaicing" a sequence of mages from the scope. This larger image is referred to as the "panorama" (see Figure 1). In the "separate" dual-view display, the panorama and the camera view are provided side by side (shown in Figure 2). In the "connected" dual-view display, the panorama and camera views are still in separate windows, with the approximate location of the camera view shown as a circular, highlighted area against the windows, but now they are visually tethered by added contours (shown in Figure 3).

The second dual-view display prototype consists of a solid and two mesh-frame representations of a 3D (m-rep)
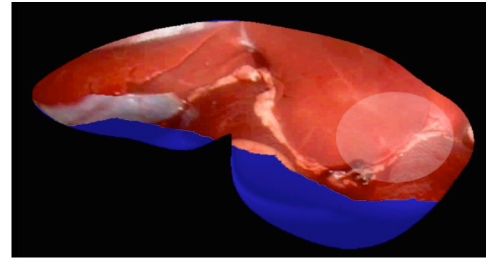

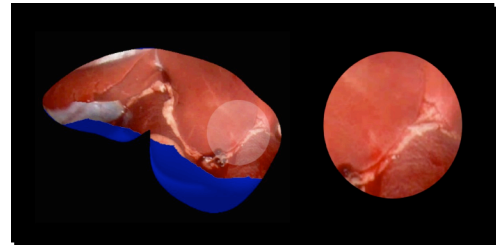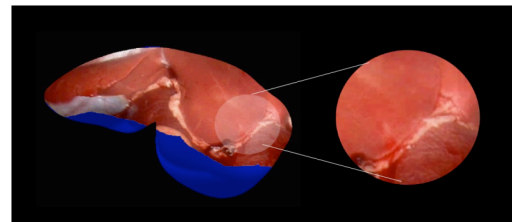Figure 1. Integrated view.

Figure 2. Separate view.

Figure 3. Connected view.

model with a tumor (shown in Figure 4). Each representation includes a global view created via nonphotorealistic rendering (left window) and a zoom-in view of the tumor in respect to the 3D model on the right window. Observation of different angles of the model or the tumor inside the modeled organ is obtainable by manipulating a mouse.

Although the concept of augmented reality for laparoscopic surgery has been developed and widely discussed since the late 90s (Azuma, 1997; Fuchs, et al., 1998; Freysinger, et al. 1997; Konishi, et al., 2007; Martin, 2005; State, et al., 2001), image-guided solutions to challenges like deformation information for operating organs or soft tissue laparoscopically are yet to be realized. The dual view display prototypes advance the research efforts of early augmentation approaches such as head mounted display systems (e.g. Fuchs, et al 1998) and optical infrared tracking technology (e.g., Konen, Scholz, & Tombrock, 1998) toward developing a user-centered system to provide seamless supports to laparoscopic surgery.

METHOD

Participants

Two residents, two fellows, and three laparoscopic surgeons (6 males; 1 female; 25 – 52 years of age) were recruited from two teaching hospitals in different states. Two colleagues (2 males) of the surgical fellows also participated
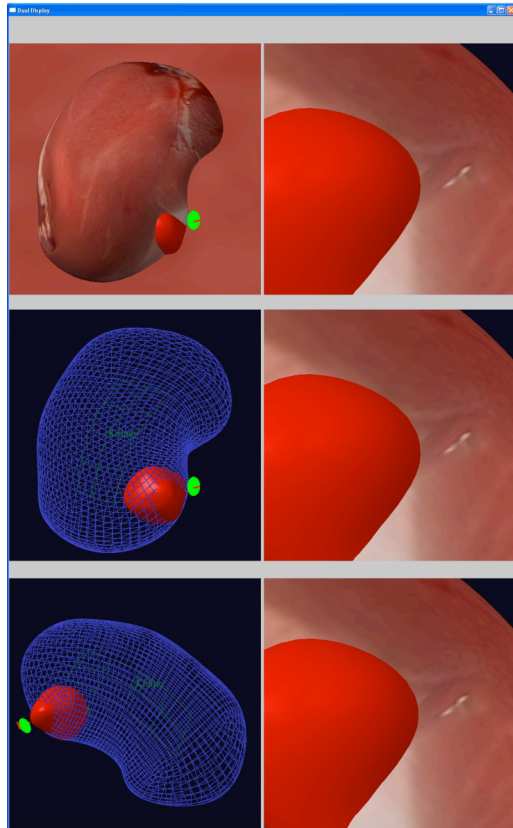
Figure 4. Solid representation of the 3-D shaped model of a kidney (top left) and the wire-frame representations of the model with a simulated tumor within (Center & bottom left). The windows on the right show the views of the 3D shaped model from the tumor.

in the interview sessions but did not complete the demographic questionnaires. All participants were in general surgery except for two who specialized in urology. The recruitment was based on a combination of chance availability and scheduled appointments. None of the participants received payments or other rewards.

Equipment

A Dell Latitude laptop computer, conference room projection screen, an audio recorder, and audiocassette tapes were used for all interview sessions. Exceptions included interviewing one participant without the projection screen, and another with projection on a small screen. Larger projections were preferred for the sessions in order to give participants some sense of immersion similar to that which should be obtained in a full prototype.

Interviews

After completing the demographic form, a visualization researcher oriented the informants to the "dual-view" display prototypes. They were asked to view the

"Integrated", "Separate" and "Connected" views while verbalizing thoughts that came to mind. Probes and interview question guides were used to direct the dialogues. The informants were then oriented to the mesh-frame display and rotated views, and were asked to comment on the display's usefulness, applications, and future development. On average, each session lasted about 30-40 minutes, and all sessions were audio recorded with permission from the informants. It should be noted that the interview procedures varied slightly among informants depending on their questions about the prototypes and their interest in elaborating on certain aspects of the interview topics.

Data analysis

Seven themes were central to the semi- structured interviews on the dual-view display prototypes: 1) ease of interpretation, 2) usefulness in individual practice, 3) potential applications, 4) identified problems, 5) improvement suggestions, 6) issues of concern, and 7) potential control devices. These themes guided the analysis of the interview transcript, which was submitted to Weft, an open-source qualitative data analysis program used for classifying and categorizing. Meaning condensation (Kvale & Brinkmann, 2009) was then used to abridge the categories for iterative analyses and interpretations.

RESULTS AND DISCUSSION

Data analysis results were summarized into four categories: 1) relevance to the informants, 2) specific applications, 3) prototypes' significance, and 4) improvements.

Relevance to the Informants

Among the separate view, the integrated view, and connected view, six informants preferred the separate view and three preferred the connected view. None considered the integrated view useful because they all perceived difficulty in sorting out the information in the embedded camera view and the global view. Also, they found that the concept of the integrated image conflicts with their medical training that emphasizes "never to get so focused on just one thing during the surgery. You want to get the big picture." This argument is particularly fascinating because it reveals a misconception that might lead to a fundamental bias against the use of the integrated display. The integrated display, in fact, will most likely lead to the easiest integration of the local information into the "big picture" so that the viewer can have two things in mind while viewing only one object (Proximity Compatibility Principle; Wickens and Carswell, 1995).

Overall, informants in general surgery found the display prototypes less applicable in their practices than the urologists. They believed that conventional CT scans serve their general needs. The general surgery practitioners also found the display impossible to use for operations on non-fixed anatomical systems such as the GI track. This may

reflect their belief that the imagery will not reflect actual deformations created by surgical manipulation. Here again we see a tendency to mistrust the accuracy of the global view, perhaps because of its de-cluttered (nonphotorealistic) rendering.

Specific Applications

All 3 surgeons and one fellow readily identified the educational value of the dual-view display. For the rest of the informants, they feared that the displays might create an additional or unnecessary learning event. Comments such as "People would be training out," "No different than using anatomical landmarks," and "Learning curve…people are not going to be familiar with this and not going to be able to use it" were some of their sentiments toward introducing the dual-view display into a training program.

However, all informants pointed out the potential of using the mesh frame display for preoperative imaging purposes to guide the examination of tumors in an organ, or "to see through things" as described by one of the residents. One resident also mentioned the benefit of having depth and distance information in the 3D dual-view display that 2D scans/images do not afford.

Other suggested applications of the dual-view display included prostatectomy, partial prostatectomy, adrenal surgery, surgical extraction or insertion, tumor ablation, gallbladder surgeries, partial nephrectomy and other kidney procedures, spleen surgery, and examination of tumor in the renal hilum.

Prototypes' Significance

Those who saw the educational value of the dual-view display believed the display would help to "maintain or accelerate the development of the training of mental concept." For instance, the display would be useful to train 1) anatomy and camera orientation, 2) mental ability to "envision the setup," and 3) knowing "what is around and what is nearby." To illustrate, the display would be helpful to show the anatomic relationship to residents, who understand the concept of a sac in gallbladder under the surface of the liver, but are uncertain of "where to go and what to do" when they slightly lift up the gallbladder and deform it from its anatomic structure.

The surgeons also contended that the dual-view display would be useful for pre-, peri- and even intra-operative planning purposes. For example, it would help to locate: 1) embedded tumor not visible on the organ surface; 2) the artery veins, collecting systems, and tumor for partial prostatectomy; and 3) the neuro-vascular bundles for prostatectomy, etc.

In term of operation, the urology surgeon saw the benefit of the display as a quick reference of "knowing where you are", and for exploring visually inaccessible areas such as the anatomy that extends back behind the bladder when looking at the prostate." Although CT scans offer similar function, they do not provide the images in real time.

Improvements

The informants raised several concerns regarding the dual-view display. Their primary concern was the accuracy of the camera registration. One surgeon emphasized that a registration tolerance of one mm or less is necessary for the display to be useful for surgical purposes. Other concerns included the accuracy of the model to represent deformity, to change the surface contour, and to reflect a growing tumor so the surgeon will make the proper incision.

For improvement and further development, the informants offered several suggestions to integrate various elements to guide the surgeon. The list included: 1) a legend and highlight of the camera view in relation to the global view of the 3 modes (separated, integrated, and connected), 2) a you-are-here map, 3) directionality and orientation information, 4) some type of translucent frame to replace the mesh frame; and 5) fewer degrees of rotation, possibly 5 degrees to 175 degrees.

In addition, the informants offered ideas on possible control devices for navigating the display views. There was no consensus among the options of a foot pedal, voice activated device, a sterile pen, or a sterile touch pad.

Conclusions

The present study describes using the expert knowledge elicitation methods of semi-structured interview and prototyping to obtain feedback from laparoscopic surgical experts during the early development of the dual-view display. Despite the limitations of the relatively brief semi-structured interviews, limited number of recruited experts, and possible bias in data analyses and interpretations, these techniques showed efficiency in generating rich and insightful data from our highly time-stressed users. We believe that expert knowledge elicitation is suitable for evaluating the early development of augmented reality display systems for surgical and training applications, and that it may have promise for the development of augmented and virtual environments more genially (Figure 5). We recommend utilizing these elicitation techniques iteratively throughout different stages of the technologies' development together with quantitative measures like mental workload assessments and performance metrics during the later phase of system development to ensure their usefulness in supporting surgical performance.

Finally, we suggest recruiting surgeons and attending surgeons for time constrained rapid prototyping. This is because an emerging theme in this study revealed that surgeons were more able to articulate their knowledge in a logical sequence than the surgical residents and even some surgical fellows. They responded methodologically with confidence. They also tended to explain and describe with scenarios for clarity. Most interesting of all, the surgeons were extremely skillful in guiding understanding of surgical procedures with cognitive walkthroughs. More investigations will be necessary on eliciting knowledge from laparoscopic experts who have different years of practices and specialty

interests. Results will enable human factors researchers and engineers to become more strategic in their recruitment of surgical experts for their investigations.
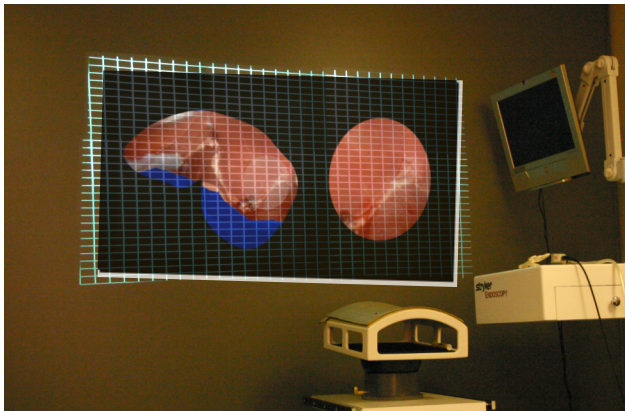

Figure 5. Future immersive dual-view display setup. The grids show the calibration of multi-projectors.

REFERENCES

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 355-385.

Carswell, C.M., Clarke, D., & Seales, W.B. (2005). Assessing mental workload during laparoscopic surgery. *Surgical Innovation, 12* (1), 80-90.

Cooke, N.J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41,* 801-849.

Felsher, J.J., Olesevich, M., Farres, H., Rosen, M., Fanning, A., Dunkin, B.J., et al. (2005). Validation of a flexible endoscopy simulator. *American Journal of Surgery, 189*(4)*,* 497-500.

Freysinger, W., Gunkel, A.R., and Thumfart, W. F. (1997). Image-guided endoscopic ENT surgery. *European Archives of Otorhinolaryngology*, *254*(7)*,* 343–346.

Fuchs, H., Livingston, M.A., Raskar.,R., Colucci, D., Keller, K., State, A., et al. (1998). Augmented reality visualization for laparoscopic surgery. *MICCAI,* 934-943.

Horn, W. (1989). MESICAR: A medical expert system integrating causal and associative reasoning. *Applied Artificial Intelligence,* A special issue on casual modeling, *3* (2), 305-336

Jetter, A., Schroder, H.H., Kraaijenbrink, J., & Wijbhoven, F. (2006). Elicitation – Extracting knowledge from experts. *Knowledge integration: The practice of knowledge management in small and medium enterprises (*pp. 65-76). Germany: Physica-Verlag Heidelberg.

Konen, W., Scholz, M., Tombrock, S. (1998). The vr project: Endoscopic image processing for neurosurgery. *Computer Aided Surgery, 3*, 144-148.

Konishi, K., Nakamoto, M., Kakeji, Y., Tanoue, K., Kawanaka, H., Yamaguchi, S, et al. (2007). A real-time navigation system for laparoscopic surgery based on three-dimensional ultrasound using magneto-optic hybrid tracking configuration. *International Journal of Computer Assisted Radiology and Surgery, 2*(1), 1-10.

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative Research interviewing* (2nd ed). Los Angeles: Sage.

Martin, R.C.G. (2005). Intraoperative magnetic resonance imaging ablation of hepatic tumors. *The American Journal of Surgery, 189,* 338-394.

Nguan, C., Girvan, A., & Luke, P.P. (2008). Robotic surgery versus laparoscopy: A comparison between two robotic systems and laparoscopy. *Journal of Robotic Surgery, 1,* 263-268.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. Amsterdam, The Netherlands: North-Holland.

Shadbolt, N., & Burton, M. (1995). Knowledge elicitation: A systematic approach. In J.R. Wilson & E.N. Corlett. *Evaluation of human work (*2nd ed., pp. 406-440). Philadelphia, PA: Taylor & Francis LTD.

Shah, S.G.S., & Robinson, I. (2006). User involvement in healthcare technology development and assessment: Structured literature review. *The International Journal of Health Care Quality Assurance, 19* (6), 500-515.

State, A., Ackerman, J., Hiota, G., Lee, J., and Fuchs, H. (2001). Dynamic Virtual convergence for video see-through head-mounted displays: Maintaining maximum stereo Overlap throughout a close-range work space. *ISAR*, 137–146.

Stefanidis, D., Haluck, R., Pham, T., Dunne, J.B., Reinke, T., Markley, S., et al. (2007). Construct and face validity and task workload for laparoscopic camera navigation: Virtual reality versus video trainer systems at the SAGES Learning Center. *Surgical Endoscopy, 21,* 1158-1164.

Swanstrom, L.L., Whiteford, M., & Khajanchee, Y. (2008). Developing essential tools to enable transgastric surgery. *Surgical Endoscopy, 22*, 600-604.

Vennix, J.A.M., Anderson, D.F., Richardson, G.P., & Rohrbaugh, J. (1992). Model building for group decision support: Issues and alternatives in knowledge elicitation. *European Journal of Operational Research, 59*(1), 28-41.

Vennix, J.A.M., & Gubbels, J.W. (1992). Knowledge elicitation in conceptual model building: A case study in modeling a regional Dutch Health care system. *European Journal of Operational Research, 59*(1), 85-101.

Wicken, C.D. & Carswell, C.M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(3), 473-494.

REAL-TIME LIGHT FALL-OFF STEREO

Miao Liao, Liang Wang, Ruigang Yang

Minglun Gong

University of Kentucky
{mliao3,lwangd,ryang}@cs.uky.edu

Memorial University of Newfoundland
gongml@gmail.com

ABSTRACT

We present a *real-time* depth recovery system using Light Fall-off Stereo (LFS). Our system contains two co-axial point light sources (LEDs) synchronized with a video camera. The video camera captures the scene under these two LEDs in complementary states(e.g., one on, one off). Based on the inverse square law for light intensity, the depth can be directly solved using the pixel ratio from two consecutive frames. We demonstrate the effectiveness of our approach with a number of real world scenes. Quantitative evaluation shows that our system compares favorably to other commercial real-time 3D range sensors, particularly in textured areas. We believe our system offers a low-cost high-resolution alternative for depth sensing under controlled lighting.

1. INTRODUCTION

Many applications, such as robot navigation and augmented reality, require real-time range information in a dynamic environment. In this paper we developed a novel system that uses the inverse square law for light intensity to estimate depth information. Based on the formulation in [1] our system uses a single camera to capture a scene under two different lighting conditions: one illuminated by a near point light source and the other by a far one. Per-pixel depth is solved based on the pixel intensity ratio and the distance between the two lights, without the need for matching pixels.

The main contribution of this paper is a novel depth range system that can generate a VGA (640×480) resolution depth map at 30Hz. Quantitative accuracy evaluation shows that our system compares favorably to other commercial 3D range sensors, particularly in textured areas. In addition, our system is made of commodity off-the-shelf components, offering an inexpensive solution to real-time, high-resolution, video-rate range sensing.

1.1. Related work

Recovering 3D shapes from images is one of the fundamental tasks in computer vision. While there is a plethora of techniques to achieve this, we will focus on the methods that are capable of generating real-time depth maps with live input.

The most common way of computing depth map is to use stereovision. Recently, several stereo methods have been developed to exploit the processing power of modern graphics hardware [2, 3, 4, 5]. Although tremendous progress has been made in stereovision, the fundamental correspondences problem remains difficult in real-world applications.

The correspondence problem can be greatly simplified with active illumination. Many real-time structured light scanners (e.g. [6, 7, 8]) can obtain high quality results. These systems typically require multiple frames, which limit the object motion, and have difficulty with high-frequency textures.

New range sensors have also been developed using shuttered light-pulse (SLP) technologies [9]. 3DV Systems, Ltd. and Canesta, Inc. [10, 11] have both developed SLP technologies. However They are either very expensive (e.g. over fifty thousand US dollars for a 3DV system) or have limited resolutions (e.g., 64×64 for a Canesta sensor).

Our system builds on the algorithms described in [1] which use the inverse-square law to recover 3D shape information. Compared to previously developed techniques, our approach only requires two images and the use of commodity off-the-shelf components provides an inexpensive way to produce high-resolution depth maps. More importantly, experiments show that our system provides better depth maps that are independent of scene texture.

2. METHODS

2.1. Light Fall-off Stereo

It is well known that the intensity of light emitted from a source of constant intrinsic luminosity falls off as the square of the distance from the object. Under this *inverse square* law, the observed intensity of a surface point p can be formulated as:

$$I_p = \frac{L(\theta)}{r_p^2} \rho(\theta, \phi), \tag{1}$$

where $L(\theta)$ is the light radiance along incident direction θ. r_p is the distance between the light source and p. $\rho(\theta, \phi)$ is the BRDF (Bidirectional Reflectance Distribution Function) of surface point p and ϕ is the viewing direction.

Now if the light source is moved away from point p along the direction θ by amount Δr. The observed intensity of surface point p under the new setting becomes:

$$I'_p = \frac{L(\theta)}{(r_p + \Delta r)^2} \rho(\theta, \phi) \qquad (2)$$

Computing the ratio of the above two equations makes the ratio between I_p and I'_p related only to the depth:

$$\frac{I_p}{I'_p} = \frac{(r_p + \Delta r)^2}{r_p^2} \implies r_p = \frac{\Delta r}{\sqrt{I_p/I'_p} - 1} \qquad (3)$$

One critical requirement for the above formulation is that the incident light direction remains the same during the illuminator's movement. An occluder is introduced in [1] to make this possible. As illustrated in Figure 1, an opaque board with a small aperture in the center is placed at the near lighting position. The plane on which the occluder lies is referred as near lighting plane. The point light source is first placed at the position of aperture, S, and illuminates the entire scene. Then the light source is moved onto a second lighting plane, which is parallel to the occluder and referred as far lighting plane. The light translates on the second plane and illuminates part of the scene each time. Consider an arbitrary point on the surface (e.g., q), it was illuminated first by a point light source at S, then by a point light source at T' that goes through S. Therefore its incident light direction remains unchanged and Equation 3 can be applied to estimate the range of q.
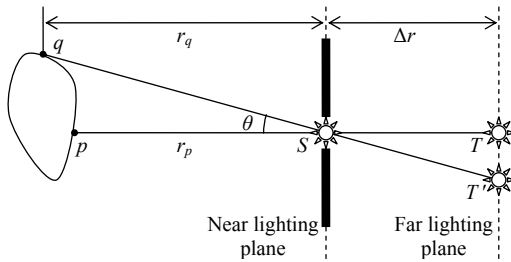
Fig. 1. The setup for recovering the depth map of a scene.

It is further argued in [1] that when the scale of the objects is much smaller than the distance of the light, the variation in incident lighting direction can be ignored. That is, the occluder can be removed and one can approximate the illumination effect obtained at position T' using the one obtained at position T. Our real-time LFS system adopts the same approximation. From the image pair captured under lighting positions S and T the per-pixel depth value can be recovered.
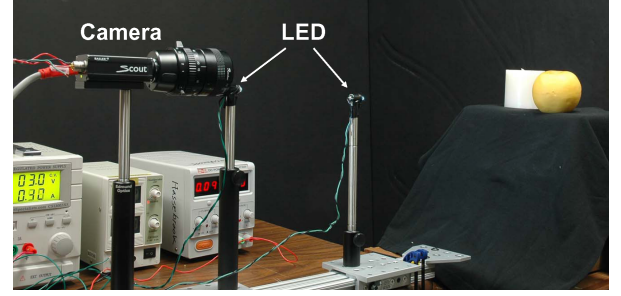
Fig. 2. Our experimental setup consists of two LED light sources and a video camera on a linear translation stage.

3. PROTOTYPE SYSTEM IMPLEMENTATION

3.1. Experimental setup

Our experimental setup consists of two 3W LED light sources and a video camera on a linear translation stage. The LEDs and video cameras are co-axial. The LEDs occlude a small part of the scene in the camera image, for which we mask out. The camera can capture 640×480 gray-scale images at 60Hz with progressive scan. The camera responds linearly to light intensity and 8-bit images are used throughout.

The two LEDs are synchronized with the camera's shutter. The camera generates a TTL signal when it opens its shutter. With each shutter pulse, the LEDs toggle their on/off state. LEDs can be switched on or off in the order of 100 nanoseconds. The LEDs heat up when powered. During the "start-up" process, the device's temperature rapidly increases, and the LED's forward current decreases until it reaches a steady state, at which point we start the capture. We have verified that the light output is very stable once the LED reaches its steady state.

While an LED is an excellent point light source, its spatial distribution of radiance is not uniform. As a result, we must radiometrically calibrate the two LEDs. The procedure is straightforward. Before running the system, we turn on the two lights in turn to illuminate a piece of white paper covering the entire field of view of the camera. The calibration object is captured by the camera under the two lighting conditions. From the two images I_n and I_f, we measure the ratio (R) of the corresponding pixel values, that is

$$R(u,v) = L_n(\theta)/L_f(\theta) = (d_n/d_f)^2 * I_n(u,v)/I_f(u,v), \qquad (4)$$

where u, v are the pixel coordinates, d_n, d_f are distance between calibration object and the near and far light respectively. Referring to equation 1, under our assumption that the incident lighting direction change is small enough, intensity variation cannot be explained by the inverse square law is attributed to the light radiance function.

3.2. Run-time algorithm

The run-time system consists of two parallel threads, one is for image capture and the other is for depth computation and display.

The capture thread waits for camera images transferred via the IEEE1394 bus and alternatively stores them into the near and the far image buffers.

In the depth computation thread, image(I_f) from far light is first corrected by calibration ratio R,

$$I_c(u, v) = I_f(u, v) * R(u, v), \qquad (5)$$

where I_c is corrected image. Then we plug $I_p = I_n(u, v)$ and $I'_p = I_c(u, v)$ into equation 3 to compute the depth of pixel (u, v).

Before the computation, we exclude those pixels that will potentially give bad results. Those pixels include saturated ones in either the near image or the far image. Saturated pixels (especially highlight areas) are usually not real measurements of light intensity; they are likely to result in inaccurate depth estimates. The pixels with intensities below a certain threshold are also excluded, because these pixels are either background or reside in shadow areas. Furthermore, low intensity values are more sensitive to noise. For those bad pixels, we simply set them to black in the depth map. This is why black holes are occasionally present in the depth map, likely the results of surface highlights and shadows. Finally, the depth map is smoothed by a mean filter.

Since graphics cards are excellent for parallel image processing, the entire depth computation pipeline is implemented on the graphics processing unit (GPU). In this case, the captured images are directly transferred to two textures on the graphics board.

4. EXPERIMENT AND RESULTS

In our experimental setup, the typical distance between near and far lights is 85mm. The valid working volume is determined by the dynamic range of the camera. With 8-bit images, it is approximately 365mm-1000mm (distance to the near light) with a depth resolution of about 4mm.

4.1. Quantitative evaluation

We first evaluate the quality performance of our system by comparing it with two other commercially available live range sensors. The first one is Canesta range sensor which is able to generate low resolution (64×64) range maps at video frame rate. The other is the Z-mini from 3DV Systems, Ltd. that can provides high resolution (maximum 640×480) live range maps.

As shown in figure 3 the target object in the first experiment (first row) is a piece of white paper glued onto a planar surface. This paper can be regarded as a perfect Lambertian reflector with constant surface albedo. In the second experiment (second row) the white paper is replaced by a piece of paper containing rich textures. The object is carefully placed so visually the principle axis of these sensors are perpendicular to the plane and go though the plane's geometric center.

One thing worth noting is that these three sensors have different fields of view, resolutions and their recovered range maps are not within the same coordinate system. These limitations make a metric comparison with ground truth difficult. To warrant a fair evaluation, we first normalize their output depth values to a uniform space. It is done by calculating a scale factor so that the sample mean of the depth values is normalized to 0.5. Afterwards we apply a plane fitting algorithm to each sample data and compute their mean square deviation. Clearly smaller variance implies better reconstruction quality.

The recovered 3D shapes and error rates of these sensors are presented in Figure 3 and Table 1 respectively. Our raw range map is processed with a 5×5 smoothing filter to reduce high frequency noise resulting primarily from the CCD camera. In general, when the sample target is textureless, all three sensors yield satisfactory results (the Canesta senor returns a single-colored depth map). However, when the target contains non-uniform surface albedo, our system outperforms the other two. It is not surprising given the fact that our depth values are recovered from the ratio of two images, effectively cancel out the surface albedo's influence. Conversely, the reconstruction model adopted by SLP sensors suffers from bias as a function of object intensity [12].

	Canesta	Z-mini	LFS
newspaper	0.0741	0.0260	0.0101
white paper	0	0.0166	0.0021

**Table 1**. Numerical errors of depth recovered by different range sensors.

4.2. Live system

Figure 4 shows some live images from our system. Shadow areas are automatically detected and masked out during the depth map computation process. The scene contains objects with different shapes and reflectance properties. The resulted depth map is fairly accurate despite the slightly non-Lambertian surface reflectance.

Our camera captures at 60fps and the off-line depth computation can achieve 60fps on a Geforce 8800 graphics card from NVIDIA. But given that two images are required to generate one depth map, our system's overall speed performance is 30fps.

5. CONCLUSION

In this paper we presented a novel system that can generate real-time depth maps. Our system, based on the formulation
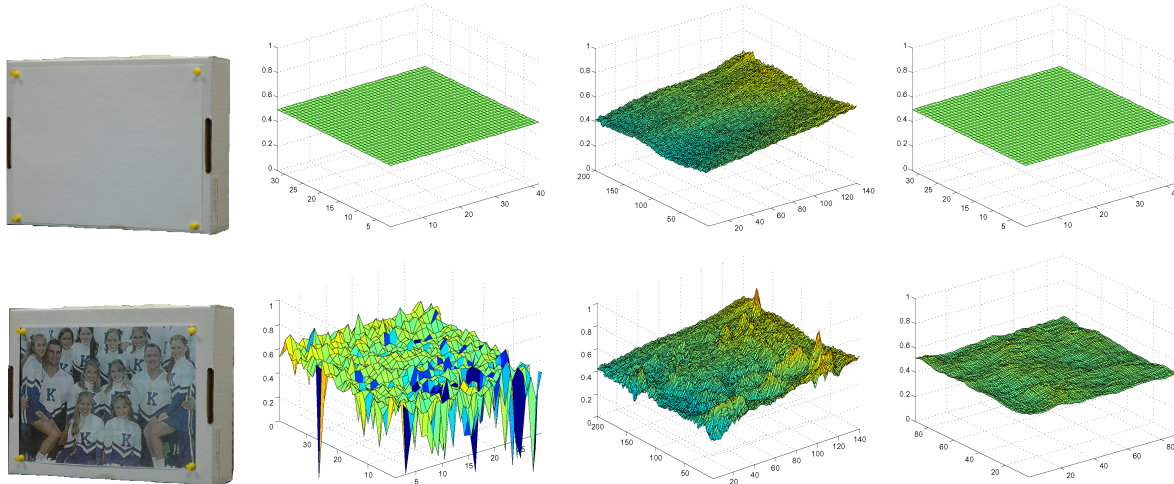
Fig. 3. Depth recovered by different sensors. From left to right: sample scenes, 3D plots of the recovered scene depth from Canesta, Z-mini and LFS. The mean depth is normalized to 0.5.
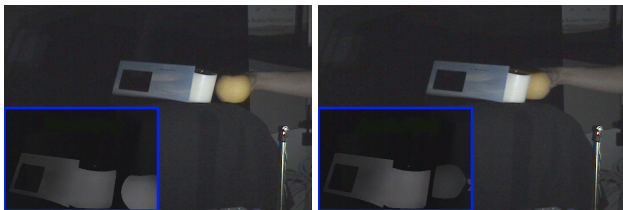
Fig. 4. Some snapshots of our real-time system results. The insets show the depth maps.

in [1], takes two images under different lighting conditions to estimate the range for each pixel, without the need for matching. Compared to commercial 3D range sensors, it is more robust to textured areas (when the object remains static or the object motion between two frames is smaller than one pixel). Our prototype is made from off-the-shelf, low cost components with a simple computation model. It can be used in low-cost embedded systems. We believe our system provides a viable alternative for 3D range sensing under controlled lighting.

6. REFERENCES

[1] M. Liao, L. Wang, R. Yang, and M. Gong, "Light fall-off stereo," *Proceedings of CVPR*, 2007.

[2] R. Yang and M. Pollefeys, "Multi-resolution real-time stereo on commodity graphics hardware," *Proceedings of CVPR*, pp. 211–218, 2003.

[3] C. Zach, A. Klaus, and K. Karner, "Accurate dense stereo reconstruction using graphics hardware," *Proceedings of EUROGRAPHICS*, pp. 227–234, 2003.

[4] N. Cornells and L. Van Gool, "Real-time connectivity constrained depth map computation using programmable graphics hardware," *Proceedings of CVPR*, pp. 1099–1104, 2005.

[5] Minglun Gong and Yee-Hong Yang, "Near real-time reliable stereo matching using programmable graphics hardware," *Proceedings of CVPR*, pp. 924–931, 2005.

[6] Olaf Hall-Holt and Szymon Rusinkiewicz, "Stripe boundary codes for real-time structured-light range scanning of moving objects," *Proceedings of ICCV*, 2001.

[7] L. Zhang, B.Curless, and S.Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," *Proceedings of 3DPVT*, 2002.

[8] Thomas P. Koninckx and Luc Van Gool, "Real-time range acquisition by adaptive structured light," *PAMI*, vol. 28, pp. 432–445, 2006.

[9] H. Gonzalez-Banos and J. Davis, "Computing depth under ambient illumination using multi-shuttered light," *Proceedings of CVPR*, pp. 234–241, 2004.

[10] G. Yahav and G. Iddan, "Optimal ranging camera," *United States Patent; no. US 6,057,909*, May 2002.

[11] C. Bamji, "Cmos-compatible 3-dim. image sensor ic," *United States Patent, no. US 6,323,942 B1*, November 2001.

[12] J. Davis and H. Gonzalez-Banos, "Enhanced shape recovery with shuttered pulses of light," *Proceedings of IEEE International Workshop on Projector-Camera Systems*, 2003.

Model Completion via Deformation Cloning Based on an Explicit Global Deformation Model[*]

Qiong Han[1], Stephen E. Strup[2], Melody C. Carswell[1], Duncan Clarke[1], and Williams B. Seales[1]

[1] Center for Visualization and Virtual Environments, University of Kentucky, USA
[2] Division of Urology, Department of Surgery, University of Kentucky, USA
qhan@engr.uky.edu

Abstract. Our main focus is the registration and visualization of a pre-built 3D model from preoperative images to the camera view of a minimally invasive surgery (MIS). Accurate estimation of soft-tissue deformations is key to the success of such a registration. This paper proposes an explicit statistical model to represent global non-rigid deformations. The deformation model built from a reference object is cloned to a target object to guide the registration of the pre-built model, which completes the deformed target object when only a part of the object is naturally visible in the camera view. The registered target model is then used to estimate deformations of its substructures. Our method requires a small number of landmarks to be reconstructed from the camera view. The registration is driven by a small set of parameters, making it suitable for real-time visualization.

1 Introduction

The distinct advantage of minimally invasive surgery (MIS) is that it induces less trauma to patients. Preoperative images reveal important substructures of target objects, which are unfortunately not visible under a laparoscopic camera view. Incorporating preoperative images into MIS is thus focused by many researchers. Among different approaches, reconstructing 3D points from a camera video sequence and registering a pre-built 3D model to the reconstructed 3D points has the strength of converting the 3D-to-2D registration to a 3D-to-3D registration.

Devernay et al. proposed a 5 step method for augmented reality of cardiac MIS [1]. [2] uses stereo images to reconstruct dense depth cues of surgical scenes. [3] fused stereo depth cues with monocular depth cues based on surface shading. Stereo based methods in general require repeatable tracking of a large number of feature points in order to reconstruct a dense set of surface points. Structure from motion (SFM) method is also adapted to MIS, and Hu et al. used a Competitive

---

Evolutionary Agent-based (CEA) method to deal with the missing data problem in SFM [4].

Statistical models of deformations and motions have also been proposed. In [5], a statistical deformation model is built from simulated finite element model (FEM) deformations of the prostate. However, proper tissue property parameters are difficult to determine for FEM simulations. [6] explicitly included material properties in their FEM simulations and built a statistical motion model to guide the deformation estimation of the prostate. [7] proposed an explicit 1D motion model to represent and compensate the motion of the mitral valve annulus.

Our driving clinical applications are laparoscopic cryoablation and laparoscopic partial nephrectomy on small renal tumors. 3D visualization of a kidney and its tumor is expected to increase the positioning accuracy of the tumor. Furthermore, surgical plans based on the same preoperative scans, from which the 3D model is built, can be visualized in real time to improve the precision of needle insertion for cryoablation or of incision site and depth for partial nephrectomy. The challenge is that there are always non-rigid intra-object deformations between the kidney in the CT scans and the kidney during an MIS. This paper proposes an explicit global deformation model, which is statistically built from a reference object and its deformed shapes. Furthermore, training data to learn a deformation model is sometimes difficult to acquire. Therefore, we propose to *clone* a learned deformation model to a new target object to guide the registration of the target object into the camera view, based on a small number of landmarks reconstructed from a camera video sequence.

Next section details our proposed method by its main steps. Section 3 describes the evaluation of the proposed method and shows the results. Section 4 concludes the paper with discussions.

2   Method

Our method takes a 5-step process shown as follows and detailed in the following subsections.

1. Build a statistical deformation model from a reference object and its deformed shapes;
2. Build a 3D model of a target object from the pre-operative computed-tomography (CT) scans;
3. Capture a video sequence of the exposed target object with a calibrated laparoscopic camera, and reconstruct 3D landmarks of the target object using the SFM method;
4. Clone the statistical deformation model to the target object model to register the target model to the reconstructed 3D landmarks;
5. Apply the deformation of the registered target object to its substructures.

2.1   A Statistical Deformation Model

The *discrete m-rep* [8] is chosen as the shape model because of its unique property of modeling and parameterizing both the surface and the interior volume of

an object. A discrete m-rep $\mathbf{M}$ consists of a quad-mesh of n_M *medial atoms* $\{\mathbf{m}_i, i = 1, 2, ..., n_M\}$. Each internal atom $\mathbf{m}_i$ has a hub position $\mathbf{p}_i$, two spokes $\mathbf{S}_i^{+1,-1}$ with a radius r_i and direction $\mathbf{U}_i^{+1,-1}$. Atoms at the edge of the quad-mesh are treated differently. For simplicity, all medial atoms are considered the same in this paper.

To learn a statistical deformation model based on m-reps, a series of deformed shapes of a reference object are captured either by a set of CT scans or by a series 3D reconstructed meshes. This paper uses the latter. An m-rep is fitted to each mesh to form a training set of m-reps. Principal geodesic analysis (PGA) [9] is applied to the training set of m-reps to form a statistical deformation model, given as a Frechét mean $\overline{\mathbf{M}}$, the first n_{PGA} principal geodesic directions $\mathbf{v}_j, j = 1, 2, ..., n_{PGA}$, representing more than 95% of the total deformation variations, and the corresponding variances λ_j of the principal geodesic directions.

Now given a set of principal geodesic components $c_j \in \mathbb{R}, j = 1, 2, ..., n_{PGA}$, a deformed reference object $\mathbf{M}_{PGA}$ can be reconstructed from $\overline{\mathbf{M}}$ and a tangent vector $\sum_{j=1}^{n_{PGA}} c_j \mathbf{v}_j$ via the *exponential map* [9]. The deformation between $\mathbf{M}_{PGA}$ and $\overline{\mathbf{M}}$ can be represented by the *residue* between the two m-reps [10], which is defined as the set of residues between all corresponding atom pairs $(\mathbf{m}_{PGA,i}, \overline{\mathbf{m}}_i)$.

Each medial atom is an element of a Riemannian symmetric space $\mathcal{G} = \mathbb{R}^3 \times \mathbb{R}^+ \times \mathcal{S}^2 \times \mathcal{S}^2$. The following operator defines the difference between a pair of atoms $(\mathbf{m}_{PGA,i}, \overline{\mathbf{m}}_i)$:

$$\mathbf{m}_{PGA,i} \ominus \overline{\mathbf{m}}_i = (\mathbf{p}_{PGA,i} - \overline{\mathbf{p}}_i, \frac{r_{PGA,i}}{r_i}, \mathbf{R}_{\mathbf{S}_{PGA,i}^{+1}}(\overline{\mathbf{S}}_i^{+1}), \mathbf{R}_{\mathbf{S}_{PGA,i}^{-1}}(\overline{\mathbf{S}}_i^{-1})) \quad (1)$$

where for any $\mathbf{w} = (w_1, w_2, w_3) \in \mathcal{S}^2$, $\mathbf{R}_{\mathbf{w}} \in \mathcal{SO}(3)$ is the rotation around the axis passing the origin $(0, 0, 0)$ and $(w_2, -w_1, 0)$ with the rotation angle being the geodesic distance between a chosen point $\mathbf{p}_0 = (0, 0, 1)$ and $\mathbf{w}$ on the unit sphere. Let $\Delta\mathbf{m}_i = \mathbf{m}_{PGA,i} \ominus \overline{\mathbf{m}}_i$. $\Delta\mathbf{m}_i$ is also an element of $\mathcal{G}$, and it is called the residue of $\mathbf{m}_{PGA,i}$ to $\overline{\mathbf{m}}_i$, which records the deformation of $\mathbf{m}_{PGA,i}$ relative to $\overline{\mathbf{m}}_i$'s coordinates.

The residue, i.e., the deformation, between a pair of atoms can then be cloned to a new atom via an operator $\oplus$:

$$\mathbf{m}_i \oplus \Delta\mathbf{m}_i = (\mathbf{p}_i + \Delta\mathbf{p}_i, r_i \Delta r_i, \mathbf{R}_{\mathbf{S}_i^{+1}}^{-1}(\Delta\mathbf{S}_i^{+1}), \mathbf{R}_{\mathbf{S}_i^{-1}}^{-1}(\Delta\mathbf{S}_i^{-1})) \quad (2)$$

where $\mathbf{R}_{\mathbf{w}}^{-1} \in \mathcal{SO}(3)$ is the inverse rotation of $\mathbf{R}_{\mathbf{w}}$.

Based on operators $\ominus$ and $\oplus$, the residue $\Delta\mathbf{M}$ between two m-reps $\mathbf{M}_{PGA}$ and $\overline{\mathbf{M}}$ and the deformation cloning of $\Delta\mathbf{M}$ to a target m-rep $\mathbf{M}_t$ are defined as follows:

$$\Delta\mathbf{M} = \mathbf{M}_{PGA} \ominus \overline{\mathbf{M}} = \{\Delta\mathbf{m}_i, i = 1, 2, ..., n_M\}, \quad (3)$$

$$\mathbf{M}_{deformed,t} = \mathbf{M}_t \oplus \Delta\mathbf{M} = \{\mathbf{m}_{t,i} \oplus \Delta\mathbf{m}_i, i = 1, 2, ..., n_M\}. \quad (4)$$

$\Delta\mathbf{M}$ is the explicit statistical deformation model learned from the reference object, which is a function of the principal geodesic components $\{c_j\}$.

2.2   A Pre-built 3D Model for the Target Object

The 3D m-rep model $\mathbf{M}_t$ is built from a manual segmentation of pre-operative CT scans of the target object by experts, and an m-rep is fitted into the segmentation using the binary fitting method described in [11]. An automatic segmentation tool will be highly desirable for this step.

2.3   Reconstruction of 3D Landmarks

Using the structure from motion (SFM) method, a dense set of object surface points $\mathbf{L}_{all} = \{\mathbf{l}_k, k = 1, 2, ..., N_{L_{all}}\}$ are reconstructed from a laparoscopic video sequence of the target object. A small subset of $\mathbf{L}_{all}$ are identified as a set of 6 to 9 anatomical landmarks $\mathbf{L} = \{\mathbf{l}_k, k = 1, 2, ..., n_L\}$. At the same time, an initial correspondence is established between the set of landmarks $\mathbf{L}$ and a set of surface points on the m-rep $\mathbf{M}_t$. This correspondence will, however, be automatically updated in the registration step whenever necessary, via the iterative closest point (ICP) method [12].

In order to get a robust reconstruction of the landmarks, fiducial markers can be used because of the small size of the landmark set. Although this step is not the main focus of this paper, the accuracy of the 3D reconstruction is crucial to the consequent steps. The effect of reconstruction errors on the registration step are evaluated in section 3.

2.4   Model Registration via Deformation Cloning

By cloning $\Delta\mathbf{M}$ to a target m-rep $\mathbf{M}_t$, we transfer the deformation learned from the reference object to the target object. As a result, we have a specific deformation model for the target object. An alignment step is required to properly clone a deformation to the target object. The alignment is described first, followed by a full description of the registration step.

Alignment step: in order to properly apply a deformation residue $\Delta\mathbf{M}(\{c_j\})$ to $\mathbf{M}_t$, $\mathbf{M}_t$ must be aligned to the mean reference object $\overline{\mathbf{M}}$ via a similarity transformation $\mathbf{T}_{sim} = \{\mathbf{p}_{sim} \in \mathbb{R}^3, r_{sim} \in \mathbb{R}^+, \mathbf{R}_{sim} \in \mathcal{SO}(3)\}$: $\mathbf{T}_{sim} = \arg\min_{\mathbf{T}} dis^2_{geodesic}(\mathbf{T}(\mathbf{M}_t), \overline{\mathbf{M}})$, where $dis^2_{geodesic}(\mathbf{M}_1, \mathbf{M}_2)$ is the squared geodesic distance between two m-reps $\mathbf{M}_1$ and $\mathbf{M}_2$ [9].

Let $\mathbf{M}_t^{aligned} = \mathbf{T}_{sim}(\mathbf{M}_t)$. A deformed target object with cloned deformation $\Delta\mathbf{M}$ is defined as $\mathbf{M}_{deformed,t} = \mathbf{M}_t^{aligned} \oplus \frac{\Delta\mathbf{M}}{r_{sim}}$, where r_{sim} is the scaling factor in $\mathbf{T}_{sim}$, and where $\frac{\Delta\mathbf{M}}{r_{sim}} = \{\frac{\Delta\mathbf{m}_i}{r_{sim}}\}$. $\frac{\Delta\mathbf{m}_i}{r_{sim}}$ means each $\Delta\mathbf{p}_i \in \mathbf{m}_i$ is replaced by $\frac{\Delta\mathbf{p}_i}{r_{sim}}$ because the translation component in an m-rep atom deformation is scale-dependent, but the scaling and rotational components are scale-independent.

$\mathbf{M}_{deformed,t}$ is then registered (fitted) to the set of reconstructed landmarks $\mathbf{L} = \{\mathbf{l}_k k = 1, 2, ..., n_L\}$. For each $\mathbf{l}_k$, there is a corresponding surface point $\mathbf{f}_k$ on the implied surface of $\mathbf{M}_t$. The fitting is implemented by minimizing an objective function:

$$\mathbf{M}'_t = \arg \min_{\mathbf{T}_{rigid}, \mathbf{M}_{deformed,t}} F(\mathbf{T}_{rigid}(\mathbf{M}_{deformed,t}(\{c_j, j = 1, 2, ..., n_{PGA}\}))) \quad (5)$$

where $F(\mathbf{M}'_t)$ has three components as $F(\mathbf{M}'_t) = t_1 F_{fit}(\mathbf{M}'_t) + t_2 F_{maha}(\mathbf{M}'_t) + (1 - t_1 - t_2) F_{leg}(\mathbf{M}'_t)$, with t_1, t_2, and $t_1 + t_2 \in (0, 1)$ as two tuning parameters: $F_{fit} = \Sigma_{k=1}^{n_L} (\frac{dis(\mathbf{f}_k(\mathbf{M}'_t), \mathbf{l}_k)}{r_{mean}})^2$ measures the fitting quality of the model to the set of landmarks by the Euclidean distance function dis and the geometric mean of the radii of all medial atoms r_{mean}; $F_{maha} = \Sigma_{i=1}^{n_{PGA}} (\frac{c_j}{\lambda_j})^2$ is the squared Mahalanobis distance between the current m-rep $\mathbf{M}'_t$ and the m-rep $\mathbf{M}_t^{aligned}$ without deformations, penalizing big deformations of $\mathbf{M}'_t$; $F_{leg} = \Sigma_{i=1} n_M f_{leg}(\mathbf{m}'_{t,i})$, where $\mathbf{m}'_{t,i}$ is a medial atom in $\mathbf{M}'_t$, and where f_{leg} is the illegality penalty term defined by equation (12) in [11]. This component penalizes shape illegalities, such as creasing or folding.

The overall algorithm is shown as follows:

1. Initialize $\{c_j\}$ to $\{0\}$, and calculate an initial alignment $\mathbf{T}_{rigid}$ to minimize $F(\mathbf{M}_t^{aligned})$;
2. Optimize $F(\mathbf{M}'_t)$ over $\{c_j\}$ and $\mathbf{T}_{rigid}$ via the conjugate gradient method until the objective function converges. Because of the compactness of the deformation model, n_{PGA} is usually smaller than 5, and the optimization usually converges within 30-40 sub-steps;
3. If $F_{fit}(\mathbf{M}'_t)$ is bigger than an empirically set threshold ε, an iteration of ICP is used to re-establish the correspondence between $\mathbf{M}'_t$ and the landmark set $\mathbf{L}$, and go back to step 2.

Step 3 is often not necessary if the initial correspondence between the small set of reconstructed landmarks L and the target m-rep is good. For majority of the testing cases, to be shown in next section, one iteration of the optimization of the objective function $F(\mathbf{M}'_t)$ is sufficient. However, by updating an initial correspondence that is of poor quality, the overall algorithm is more robust to correspondence errors.

2.5   Deformation Propagation to Substructures

The target models before and after the registration are used to imply a deformation field for the interior and the adjacent exterior volume of the target object. The deformation field is propagated to the substructure volume, voxel by voxel. Because of the enforced legality of the deformed $\mathbf{M}'$ by the component F_{leg}, the volumetric legalities of both the models, before and after the registration, are guaranteed. Therefore, the implied deformation field is guaranteed to be legal. Next section evaluates the proposed method.

3   Result

In order to evaluate the proposed method, a set of kidney models with synthetic deformations is generated. Synthetic data provide the ground truth to better evaluate our method. Also, the impact of reconstruction errors by the SFM method is studied. One set of *in vivo* data is also used to test our method.

The rationale of using synthetic deformations is that the types of deformations a kidney undergoes during an MIS can be well described and modeled by experienced surgeons so it is a reasonable approximation to population deformations. However, our method can be applied to dynamic CT or range data sets to learn arguably more realistic organ deformations.

3.1   Generating Synthetic Testing Data

There are two parts of data generations:

- Generation of the statistical deformation models: 20 kidney m-reps $\mathbf{M}_{kid,i}$, $i \in [1, 20]$ from different patients are used. A series of simulated deformations are applied to each kidney m-rep. Each kidney m-rep and its deformed shapes are used to build a statistical deformation model $\Delta\mathbf{M}_{kid,i}(\{c_j\})$ of the reference m-rep $\overline{\mathbf{M}}_{kid,i}$. Each statistical deformation model is then used to guide the registration of all the other 19 kidneys. In total there are 20×19 registration results. A tumor m-rep is also added to each kidney m-rep.
- Generation of video sequences for SFM reconstructions: a diffeomorphic deformation, independent from the deformations used to generate the statistical deformation models, is applied to the m-rep implied surface meshes of the kidney and tumor.

  A kidney texture image, stitched from an *in vivo* video, is used as the texture for each deformed kidney mesh $\mathbf{Mesh}_{kid,i}$. Using the parameters of a calibrated Stryker laparoscope, a series of 15 images $\mathbf{I}_i$ are generated at the resolution of 640×480 to cover about half of each kidney surface, assuming no deformations among these image frames. A set of 100 surface points are randomly selected as the ground truth reconstructed surface points $\mathbf{L}_{truth,all,i}$. 6 to 9 landmarks of anatomical significance are selected from each mesh as the set $\mathbf{L}_{truth,i}$. Initial correspondence between $\mathbf{L}_{truth,i}$ and the m-rep $\mathbf{M}_i$ is also automatically established.

3.2   Experimental Results from Synthetic Data

Guided by the statistical deformation model learned from the reference kidney $\mathbf{M}_i \in [1, 20]$, each m-rep $\mathbf{M}_j, j \neq i$ was registered into its video sequence $\mathbf{I}_j$ to acquire the registered m-rep $\mathbf{M}'_{i,j}$. Each $\mathbf{M}'_{j,i}$ was compared to the ground truth landmark points $\mathbf{L}_{truth,j}$ to calculate the average point-to-point distance (APD), and $\mathbf{M}'_{j,i}$ was also compared to each ground truth mesh $\mathbf{Mesh}_{kid,j}$ to calculate the average surface distance (ASD) Each $\mathbf{M}'_{j,i}$ was then used to estimate and apply propagation deformations to its tumor model. The deformed tumor model was compared to the ground truth tumor mesh $\mathbf{Mesh}_{tumor,j}$ to calculate the ASD. A deformation model and 3 testing kidney models are shown in figure 1.

All the experiments were conducted with different levels of Gaussian noise added to the reconstructed surface points $\mathbf{L}_{all,j}$: the standard deviations are 1, 3, and 5 voxels. The size of 1 voxel is approximately $0.78mm$. The average experimental results are shown in table 1. The deformation propagation errors of the tumor are
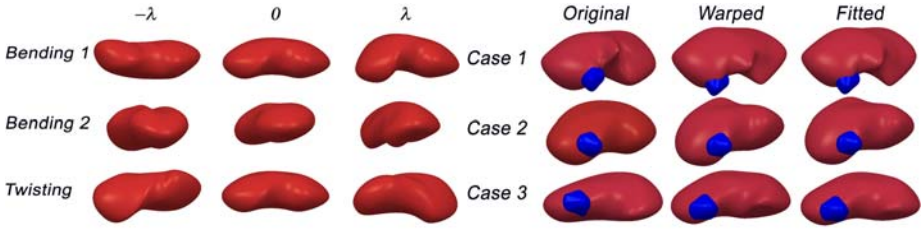
Fig. 1. Left: 3 main modes of a deformation model: each row shows one mode from $-\lambda$, 0, to λ; Right: 3 testing kidney m-reps, from left to right: the original target kidney m-rep $\mathbf{M}_i$ (in red), the ground truth surface mesh of the kidney and tumor (in blue) reconstructed from warped object volume, and the registered m-rep $\mathbf{M}'$ with the deformation applied to its attached tumor model

Table 1. All units are in voxels, with the size of $0.78mm$, except the number of iterations

STD of Noise	Kidney Avg. APD	Kidney Avg. ASD	Tumor Avg. ASD	Avg. Number of Iterations	Kidney Avg. ASD Without ICP
1	0.57 ± 0.88	0.75 ± 0.65	1.38 ± 0.69	1.36 ± 0.53	1.14 ± 0.78
3	2.24 ± 1.35	2.59 ± 0.95	3.25 ± 1.06	3.00 ± 0.87	3.76 ± 2.03
5	3.27 ± 1.45	3.41 ± 1.30	4.19 ± 1.41	5.81 ± 1.54	6.82 ± 3.44

bigger than the registration errors of kidneys, which is expected. As the noise level for the reconstruction error increases, the registration errors increase too, but at a slower pace. At a lower noise level, most registrations only require 1 iteration of optimization. However, the ICP step is necessary to keep the registration robust as the noise increases. The last column shows that the registration results deteriorate rapidly without the ICP to correct a poor initial correspondence.

3.3   Results from a Set of *in vivo* Data

A CT scan of $1mm \times 1mm \times 3mm$ was used to build the initial m-rep model for the target kidney. Because of the lack of enough training data, the deformation model built from the synthetic data was used to guide the registration of the m-rep to the video sequence. There is no ground truth surface mesh available. A dense set of 200 surface points were reconstructed from the video sequence and were used in the registration. The average distance between the surface points to the implied boundary surface of the registered m-rep is $2.65mm$.

4   Discussion

Our method has the advantages as follows: the registration via deformation cloning uses a statistical deformation model learned from often very limited training data, and the registration completes the deformed target object; only

a small number of reconstructed landmarks are required as long as a good correspondence between the landmark set and the target model is established; the registered deformations of the target object can be used to estimate deformations to important substructures.

We are working on live animal experiments to further validate our method. One challenging but rewarding extension of our method is to combine and apply multiple deformation models to a new target object.

References

1. Devernay, F., Mourgues, F., Coste-Maniere, F.: Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery. In: Proceedings of the International Workshop on Medical Imaging and Augmented Reality (MIAR 2001), pp. 16–20 (2001)
2. Lau, W., Ramey, N., Corso, J., Thakor, N., Hager, G.: Stereo-based endoscopic tracking of cardiac surface deformation. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3217, pp. 494–501. Springer, Heidelberg (2004)
3. Lo, B.P., Scarzanella, M.V., Stoyanov, D., Yang, G.Z.: Belief propagation for depth cue fusion in minimally invasive surgery. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 104–112. Springer, Heidelberg (2008)
4. Hu, M., Edwards, P., Figl, M., Hawkes, D.J.: 3D reconstruction of internal organ surfaces for minimal invasive surgery. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 68–77. Springer, Heidelberg (2007)
5. Mohamed, A., Davatzikos, C., Taylor, R.: A combined statistical and biomechanical model for estimation of intra-operative prostate deformation. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2489, pp. 452–460. Springer, Heidelberg (2002)
6. Hu, Y., Morgan, D., Ahmed, H.U., Pends, D., Sahu, M., Allen, C., Emberton, M., Hawkes, D., Barratt, D.: A statistical motion model based on biomechanical simulations for data fusion during image-guided prostate interventions. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 737–744. Springer, Heidelberg (2008)
7. Yuen, S.G., Kesner, S.B., Vasilyev, N.V., Nido, P.J.D., Howe, R.D.: 3D ultrasound-guided motion compensation system for beating heart mitral valve repair. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 711–719. Springer, Heidelberg (2008)
8. Pizer, S.M., Fletcher, T., Fridman, Y., Fritsch, D.S., Gash, A.G., Glotzer, J.M., Joshi, S., Thall, A., Tracton, G., Yushkevich, P., Chaney, E.L.: Deformable m-reps for 3d medical image segmentation. International Journal of Computer Vision - Special UNC-MIDAG issue 55(2), 85–106 (2003)
9. Fletcher, P.T., Lu, C., Pizer, S.M., Joshi, S.: Principal geodesic analysis for the nonlinear study of shape. Transactions on Medical Imaging (TMI) 23(8), 995–1005 (2004)
10. Lu, C., Pizer, S.M., Joshi, S., Jeong, J.Y.: Statistical multi-object shape models. IJCV 75(3), 387–404 (2007)
11. Han, Q., Merck, D., Levy, J., Villarruel, C., Damon, J.N., Chaney, E.L., Pizer, S.M.: Geometrically proper models in statistical training. In: Karssemeijer, N., Lelieveldt, B. (eds.) IPMI 2007. LNCS, vol. 4584, pp. 751–762. Springer, Heidelberg (2007)
12. Besl, P., Mckay, N.: A method for registration of 3-d shapes. IEEE TPAMI 14(2), 239–256 (1992)

A Research Portfolio for Innovation in the Surgical Environment

Gerald R. MOSES,[a, 1] PhD, Adrian E. PARK,[b] MD,
[a] University of Maryland Medical Center, Baltimore MD
[b] Department of Surgery, University of Maryland Medical Center

Abstract. The University of Maryland Medical Center and School of Medicine have sponsored a program of research targeted at the enabling of technologies for enhanced training, clinical effectiveness and patient safety. The pillars of this research included scientific approaches related to Informatics, Smart Image, Simulation and Ergonomics and Human Factors. The evolving research effort opened the door to a revised concept of basic surgical sciences that underpin training and performance in the operative environment.

Keywords. Surgery, training, innovation, surgical basic sciences

Background

The phrase, operating room of the future (ORF), has been used to describe the development of medical technology and the improvement of function and safety of the perioperative environment. The research program in the Department of Surgery at the University of Maryland has extended the meaning of the ORF to the study of functions and interactions of people, processes and technology producing a safe and efficient operating suite.

The Research Portfolio

For five years, the University of Maryland Medical Center and School of Medicine have sponsored a program of research targeted at the enabling of technologies for enhanced training, clinical effectiveness and patient safety. Initially, under the rubric of "The Operating Room of the Future" various pillars of research were established that proposed to advance the state of medicine, notably surgery. The pillars included scientific approaches related to Informatics, Smart Image, and Simulation. The evolving research effort opened the door to a revised concept of basic surgical sciences that underpin training and performance in the operative environment.

Developments led to two important changes; the adoption of a new mantra, Innovation in the Surgical Environment, to replace the Operating Room of the Future; and the addition of another research pillar, that of Ergonomics and Human Factors.

---
[1] Corresponding Author: Gerald Moses, MASTRI Center, Division of General Surgery, University of Maryland Medical Center, 22 S. Greene Street, Baltimore, MD 21201; gmoses@smail.umaryland.edu

Progress has been achieved in each of the pillars of research, as reported at a recent annual conference that sought to apply lessons learned form the high-stakes environments of aviation and astronautics to the practice of surgery.

Research Pillars

The medical informatics pillar includes a Perioperative Scheduling Study, a study of workflow around performance indicators in the peri-operative environment and building a graphical dashboard to allow data mining and trend analysis of operating indicators. The surgical simulation pillar entails both physical and cognitive simulation for training with emphasis upon laparoscopic surgery. A third pillar is entitled "Smart Image" in which we are seeking to push the boundaries of real time deformable image registration with a goal of performing the 1st fully smart image guided laparoscopy. A recently added pillar of ergonomics and human factors addresses the impact of stress movements and position upon the surgeon performing minimally invasive or "open" procedures.

*Informatics: Workflow and Operations Research for Quality (WORQ)*

The Perioperative Scheduling Study is looking at how using post-operative destination information during the process of surgery scheduling can influence congestion in post-operative units such as intensive care units (ICUs) and intermediate care units (IMCs), which lead to overnight boarders in the post-anesthesia care unit (PACU). We have developed a mathematical congestion evaluation model for evaluating congestion in post-operative units, including ICUs, IMCs, and floor units. This model requires data about post-operative destinations and length-of-stay distributions for different types of surgeries. We have analyzed data about cardiac surgeries from two years and have analyzed UMMC financial records for all of the surgical cases for fiscal year 2007. We have developed an algorithm for predicting bed requirements based on the surgical schedule and have conducted a preliminary study comparing these predictions to other prediction methods for two units. The preliminary results show that the new bed requirements prediction method is more accurate.

Informatics: Operating Room Glitch Analysis (OGA)

The OGA project, focusing on institutional learning, is looking at the workflow around performance indicators in the peri-operative environment and building a graphical dashboard to allow data mining and trend analysis of operating indicators.

We have integrated into the data architecture a javascript based bubble chart that provides several interactive features to allow thorough data discovery. The bubble chart can play over time to see how the size of the bubbles change, which relates to the number of cases performed, as well as their x and y axis location. The x and y axis can represent delay duration, actual procedure time, scheduled procedure time, or turnover time. The bubble can also be tagged to provide a contrail to show performance over time. Figure 1 indicates an analysis of service delay as related to average length of surgical procedure.
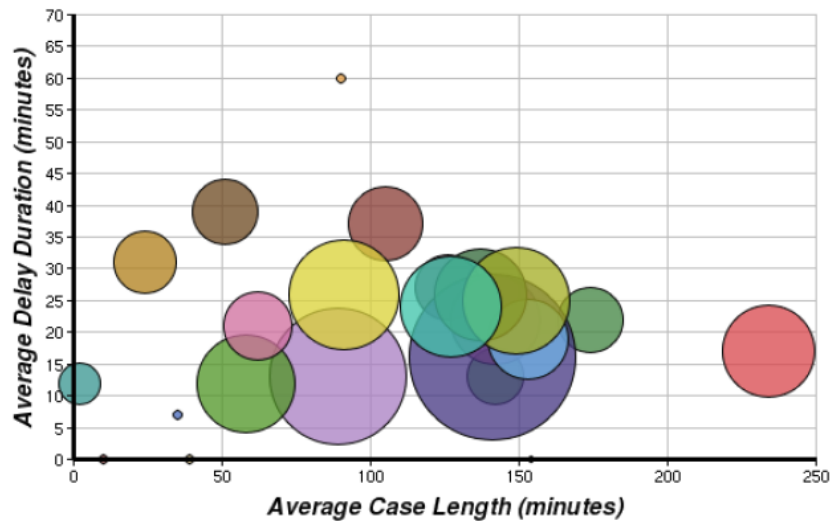
Figure 1. Service delay as related to average length of surgical procedure.

Informatics: Video Summarization of Key Events in Surgery

The technique of summarization is used when confronted with the task of gleaning succinct information from large amounts of data. For example, our national intelligence services use both machine and human analysis to prepare the daily Intelligence Summary for the President. A similar challenge is presented to those who train surgeons using a vast archive of surgical video. A key element in teaching is the extraction of the right video event to make the critical point to surgical trainees.

Recent decades have seen an increasing use of VR and simulation aids in surgical training. The typical approach is to use sensors to capture the kinematics of the tools, as well as force/torque measures. One thread of work directly analyzes these measurements to construct Markov Models that describe the state and transitions for a surgical procedure, and it is then shown that the transition probabilities between states are different at different levels of expertise.

An alternative approach is for an expert to look at the video of the surgical procedure (or training), identify key steps/events (either done well or incorrectly), and then judge the skill level of the performer. This approach can bring to bear the expert's knowledge and intuition of the complex interaction between tools, movements, organs, cutting planes etc. The drawback however is that it requires the review of a video that can be very time-consuming. We propose to address this problem by developing techniques to automatically identify key scenes/events in a video of laparoscopic surgeries.

Simulation

We are conducting multiple studies of the effects of physical box trainers, virtual reality (VR) trainers, and mixed modality training for acquiring laparoscopic surgery skills. These studies support the actions and operations of the Maryland Advanced Simulation Training, Research and Innovation (MASTRI) center. Additionally, we are developing a cognitive simulator and building knowledge representations based on ontology of focused human anatomy/physiology to emulate the surgical/clinical experience. The cognitive simulator, the Maryland Virtual Patient, has been developed by construction of a computational model of the cognitive agent, and by testing the goal- and plan-based reasoning component and its interaction with the interoceptive and language perception modules and verbal, mental and physical action simulation modules.

We have continued to work on the natural language substrate of the system, concentrating on enhancements required for processing dialog (not expository text). Further, we have implemented an enhanced microtheory of indirect speech acts, and continued working reference resolution algorithms.

The research work encompasses work targeted upon the acquisition of ontology and lexicon knowledge., and improvement of the DEKADE user interface. The current version of the cognitive simulation system includes multiple scenarios of physician-patient interface related to LERD/GERD patient conditions.

Smart Imaging

Surgical practice is considered among the most complex and difficult fields. That no two patients are exactly alike is one of the challenges that make it so. Anatomic and physiologic differences make each case unique. In surgery, these variations can complicate an operation; the discovery of unexpected anatomical variations often requires a surgeon to stray from standard, well-practiced techniques to attempt a novel approach to the procedure. With novelty comes a reduced margin of safety. This situation is exacerbated by a trend toward further physical separation between the patient and interventionalists (e.g., surgeons, endoscopists, radiologists) and a greater dependence on an image of the patient's (target) anatomy to effect therapy or establish a diagnosis.

"Smart image," as we have defined it, refers either to the process of extracting elements from an environment and imparting them to an image or to acquiring elements from within a scene and enhancing them. The result in either case is a more meaningful visualization of the operative field. Although many applications exist within this definition, Maryland's smart image team is working toward performing the first laparoscopic surgery guided completely by smart image.

Typically in laparoscopic procedures, diagnostic imaging—including x-rays, computerized tomography (CT), and magnetic resonance imaging (MRI) scans—can provide a preview of patient physiology. Often, however, these diagnostic images are in a static format that does not allow the care provider to interact meaningfully with the information the images contain. Current advances in smart imaging can be used to improve patient safety by providing the caregiver with a more interactive experience. A set of two-dimensional (2D) slices of a CT scan can be transformed into a three-

dimensional (3D) computer model so that surgeons can preview a realistic view of the patient's anatomy before an operation. This type of smart imaging provides an interactive "fly-through" view that allows the surgeon to explore the anatomy in detail.

With advances in computing power, these previews could be mapped more realistically to interactive simulators that would permit rehearsal of a surgical procedure that might include attempts at novel approaches before surgery begins. During real surgery, these smart diagnostic images could be integrated into the surgeon's actual view of the patient.

We are working toward matching the minimally invasive surgeon's video view of the surface anatomy with computer-generated models from Digital Imaging and Communications in Medicine (DICOM) data sets. Such imaging could provide the surgeon with real-time "x-ray vision" during the operation. Thus, the underlying structure, such as the position of a tumor beneath the surface of a larger anatomic structure or blood vessels within the liver, could be seen. Vessels could be contrast-enhanced in a single, high-resolution CT scan before the surgery. Then, during surgery, low-dose/low-resolution CT scans could be used to transform the high-resolution CT image to match the movement of the patient's anatomy during surgery. This would allow intraoperative visualization of anatomy that retains the enhanced contrast vessels, a unique ability that is not possible at present.

CT scans can provide enhanced intraoperative visualization of deep structures far superior to that of laparoscopes. However, the use of continuous CT exposes the patient and surgeon to a radiation level that remains a concern. Therefore, a major thrust of our work is to design, develop, and test several dose-reduction strategies and to incorporate these into our proposed continuous CT-guided surgical navigation system. Our preliminary work suggests that our strategies would allow us to lower the net radiation exposure to the patient to levels commonly viewed as safer in cardiac catheterization and interventional radiology procedures. In the long term, we also propose using telemanipulators to remove surgeons from the CT room and thereby shield them entirely from radiation exposure while they are performing the procedure.

Ergonomics and Human Factors

Recently, a fourth pillar was added to our research portfolio, that of Ergonomics and Human Factors. These are two related branches of study that examine the relationship between people and their work environment. Ergonomics often focuses on the physical environment and the human body, while human factors center more on the cognitive aspects of performance. The same ergonomics and human factors techniques credited with making industrial processes safer and more efficient can be applied to the analysis and improvement of OR operations. Tools, such as video analysis and motion tracking, can be used to analyze current practices, identify inefficiencies and dangers, develop solutions, and measure improvement. "Best practices" to maximize safety and efficiency can be developed based on empirical data.

Our discussion of workflow to this point has taken a macro or panoramic view; for example, how might we most effectively track and bring together the people and assets necessary to ensure that a patient's surgical experience is safe and efficient. Through human factors and ergonomics, we have the ability to focus on a more micro-level

analysis, such as measurements of surgeon/instruments interface and how the physical interface between the surgeon and the patient could be improved.

In the future, OR workspace layout would be optimized through ergonomic data and human factors analysis, and this optimization would lead to the establishment of "best practices" for an array of surgical operations. Proper layout would reduce risks of infection, speed operations, and reduce fatigue of surgeons and staff, all elements that could contribute to a reduction in adverse events and improved patient safety.

Future Vision of the Operating Room Environment

Well-trained care providers, who have reached a level of proficiency on realistically simulated patients, are supported by an array of smart technology enabling surgical procedures to be performed in an ever safer environment. Cases start on time with all team members informed of the goals and possible trouble spots of each operation. Contingency plans are in place for dealing with anticipated complications. The smart environment checks that all required equipment and people are present and cross-checks drugs and blood products brought into the room, ensuring patient compatibility in terms of allergies and blood type. Surgeons do not have to fight fatigue and discomfort during surgery, as the layout of the surgical workspace is ergonomically correct. Thus, the time and effort needed to perform surgery is minimized and improvement of both technique and outcomes is realized.

A New Set of Basic Surgical Sciences

The potential of surgical care in the future can be realized by incorporating into the training of surgeons a new set of basic surgical sciences, those of advanced imaging, informatics systems, simulation and ergonomics and human factors. These do not replace the well established scientific bases of anatomy, physiology, pathology and related areas of study. Rather they add a vital underpinning to the knowledge and expertise required of future practitioners.

References

[1] Shekhar, R., Dandekar, O., Kavic, S.,George, I., Mezrich, R., and Park, A. "Development of continuous CT-guided minimally invasive surgery," in Medical Imaging 2007 Visualization and Image-Guided Procedures, San Diego, CA, USA, 2007, pp65090D-8.
[2] Shetye, A. and R. Shekhar, "A statistical approach to high-quality CT reconstruction at low radiation doses for real-time guidance and navigation," in Medical Imaging 2007: Physics of Medical Imaging, San Diego, CA, USA, 2007, pp. 65105U-11.
[3] Lee G, Kavic SM, George IM, Park AE (2007) MIS surgical ergonomics: Future Trends, Annual conference of Medicine Meets Virtual Reality (MMVR), Long Beach, CA.
[4] Moses G.R., Seagull FJ, George I.M.and Park A.E. The MASTRI Center – Medical Simulation for Skill Acquisition. Proceedings of MODSIM World Conference 2007